EFFECTS OF A PEER TUTOR TRAINING PROGRAM ON TUTORS AND TUTEES WITH SEVERE DISABILITIES IN ADAPTED PHYSICAL EDUCATION

by

Drew James vonLintel

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Exercise and Sport Science

The University of Utah

December 2015

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of **Drew James vonLintel**

has been approved by the following supervisory committee members:

Hester Henderson	, Chair	**June 18, 2015** Date Approved
Elizabeth Cook	, Member	**June 19, 2015** Date Approved
Lynn Durrant	, Member	Date Approved
Arwen Fuller	, Member	**June 16, 2015** Date Approved
Breda O'Keeffe	, Member	**June 16, 2015** Date Approved

and by **Janet Shaw**, Chair/Dean of

the Department/College/School of **Exercise and Sport Science**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

This dissertation examines the efficacy of peer tutor training in adapted physical education (APE). A peer tutor evaluation form was created to assess the skills of untrained peer tutors (n = 12). Once skills were assessed, a peer tutor training protocol was created. The protocol was implemented in a peer tutor training program. After peer tutors were trained, they participated in an APE peer tutor program for students with severe disabilities in elementary school (K-6). This study measured the effects of trained and untrained peer tutors on motor performance, number of steps. Peer tutor attitudes were also evaluated. The study employed a single-subject multiple baseline design with 24 participants (12 students with severe disabilities, 12 students from general population) in a public elementary school. The acquisition of motor skills was determined using the TGMD-2; exchange of verbal information was assessed using the peer tutor evaluation form, steps were counted for the entire class time using pedometers.

Results of this study showed trained peer tutors had a positive effect on tutee motor skill acquisition as represented by statistically significant t-test results, and step counts as represented by level change in multiple baseline data. Peer tutor performance was also effected by training as illustrated by substantial level change and minimal overlap in multiple baseline data. Change in peer tutor attitude was also statistically significant as represented by a Wilcoxson signed rank test.

TABLE OF CONTENTS

ABSTRACT iii

LIST OF TABLES vi

LIST OF FIGURES vii

Chapters

1. INTRODUCTION 1

- Introduction 2
- Literature Review 6
- Statement of Problem 15
- Purpose Statement 16
- Research Questions 16
- Method 16
- Procedure 19
- Design and Analysis 21

2. STUDY 1: DEVELOPMENT AND INTITIAL VALIDATION OF THE PEER TUTOR EVALUATION INSTRUMENT 25

- Abstract 26
- Introduction 26
- Method 30
- Procedure 34
- Results 37
- Discussion 37
- Conclusions 38

3. STUDY 2: EFFECTS OF PEER TUTORTRAINING ON PEER TUTOR PERFORMANCE IN ADAPTED PHYSICAL EDUCATION 43

- Abstract 44
- Introduction 44
- Method 49
- Instrumentation 50
- Procedure 53

Design and Analyses 54
Results 57
Tutee Step Count 65
Tutee Skill Acquisition 71
Social Validity 72
Discussion 72
Conclusions 74

4. STUDY 3: EVALUATION OF PEER TUTOR ATTITUDES TOWARD STUDENTS WITH SEVERE DISABILITIES IN ADAPTED PHYSICAL EDUCATION 80

Abstract 81
Introduction 81
Method 84
Procedure 87
Design and Analyses 89
Results 91
Verbal Response 94
Qualitative Findings 94
Discussion 96
Conclusions 98

5. CONCLUSION 101

Major Findings 102
Limitations 106
Discussion 106
Recommendations 108

Appendices

A. SCORING PROTOCOL – PEER TUTOR EVALUATION INSTRUMENT 109

B. PEER TUTOR TRAINING PROTOCOL 114

C. PEER TUTOR EVALUATION – MAIN TRAINING SESSION 126

D. 3-POINT LIKERT SCALE ATTITUDE INSTRUMENT 127

REFERENCES 129

LIST OF TABLES

Tables

2.1 Content validity steps (Lynn, 1986) 40

2.2 Peer tutor evaluation instrument goals and objectives 40

2.3 Expert reviewers' ratings of content validity of items 41

3.1 Peer tutor teacher performance scores – PTEI: mean and standard deviation 76

3.2 Tutee step count – mean and standard deviation 76

3.3 TGMD-2 Mean scores for pretest and posttest including percent change and percent mastery 77

4.1 Peer tutor 3-point Likert score selections: mean and standard deviation 99

4.2 Percentage positive and negative peer tutor attitude – baseline, intervention 99

LIST OF FIGURES

Figures

2.1 Peer tutor evaluation instrument 42

3.1 Peer tutor performance scores on the Peer Tutor Evaluation Instrument (PTEI) 78

3.2 Peer tutee step counts 79

4.1: Peer tutor 3-point Likert scale attitude selections 100

CHAPTER 1

INTRODUCTION

Introduction

Children with severe disabilities (e.g., autism spectrum disorder [ASD], intellectual disability, multiple disabilities, other health impairment, traumatic brain injury, etc.) have unique needs in educational settings (Block, 2007). They have lower academic skills, deficits in social skills, and lags in gross motor skills compared to typically developing peers (Block, 2007; Dunn & Leitschuh, 2010). These students may require one-on-one instruction for academic skills and assistance with self-help procedures, making them less likely candidates for inclusion. Teacher prompting may be required for response to academic probes and social interaction with typically developing peers.

The number of students with disabilities attending our nation's public schools has grown considerably in the past decade (Cervantes, Lieberman, Magnesio, & Wood, 2013). This growth in enrollment comes at the same time as resources decrease. Budget cuts are hampering educational support for students with disabilities; the number of paraprofessional support staff has consistently been cut in the past 10 years, leading to further limited support for higher numbers of students in the classroom (Cervantes et al., 2013). Moreover, larger class sizes make it difficult to provide individualized education to each student with a disability (Cervantes et al., 2013; Snyder & Dillow, 2011).

This lack of resources puts a substantial limitation on adapted physical education (APE) for students with severe disabilities. APE is a direct service that is required by law (PL 94-142, the Education for All Handicapped Children Act, & Individuals with Disabilities Acts - IDEA, 1990, 1997, & 2004) to be provided to children with disabilities. If individually designed physical education is prescribed in a child's IEP, the

public agency responsible for the education of that child must provide the services directly or make arrangements for those services to be provided through other public or private programs (IDEA 2004). The law states that resources must be allocated for teaching APE to students with disabilities; however, these resources continue to be limited, placing a heavy burden on practitioners and sacrificing educational experiences for students.

Studies examining Academic Learning Time-Physical Education (ALT-PE) and students with disabilities in inclusionary physical education classes suggest that students with disabilities consistently have lower amounts of ALT-PE (e.g., learning time) than their typically developing peers (Knowles, Aufderheide, & McKenzie, 1982; Silverman, Dodds, Placek, Shute, & Rife, 1984). Students with severe disabilities often lose out on learning experiences and practice (e.g., ALT-PE) because they require more support (e.g., specialized instruction, and adapted activities) from the physical educator. If the physical educator is unwilling or does not have the knowledge or resources to adapt their teaching strategies and environment to provide students with severe disabilities with the necessary learning supports, these students will not benefit from being in inclusionary classes. Researchers argue that excessive class sizes and inappropriate equipment lead to the unfortunate circumstance that students with severe disabilities are often overlooked in inclusionary settings (Wiskochil, Lieberman, Houston-Wilson & Peterson, 2007). These findings suggest instructional programming can be severely lacking for students with disabilities in integrated or inclusive settings (Webster, 1987; Wiskochil et al., 2007). Hence, it is important to have an appropriate setting for these students to succeed in physical education.

Adapted physical education classes can be an effective setting for students with severe disabilities; however, there are some constraints in this setting as well. Time allocated to APE is often restricted. Since some elementary APE classes are only 30 minutes per week (Piatti, Beets, & Combs, 2009), once transition time is subtracted, learning time can be less than 25 minutes. Each additional minute of the class should be spent teaching motor and fitness skills and helping students learn to enjoy physical activity, because it is so important in their lives. Unfortunately, too much time may be spent organizing and motivating students (Dunn & Leitschuh, 2010).

The limited support and lack of time for motor skill practice in APE classrooms translates into less motor skill acquisition for students with severe disabilities (Klavina, & Block, 2008). Moreover, these students have less chance of success in the APE due to inadequate teacher support, insufficient individualized instruction, and minimal paraprofessional staff (Block, Oberweiser, & Bain, 1995). Students with severe disabilities need to experience successful motor engagement just like their typically developing peers (Block, 1994; d'Arripe-Longueville, Gernigon, Huet, Winnykamen, & Cadopi, 2002; Ensergueix & Lafont, 2010). This success may be elusive; however, a lack of support and time may be mitigated by implementing a peer tutor program.

A potential solution to improving student-learning outcomes in physical education settings is the use of peer tutors (Block, 2007; Hodge, Lieberman, & Murata, 2012). Peer-assisted learning strategies (PALS) is an evidenced-based program that includes teaching strategies using peers tutors to support direct instruction (PALS; Rosenshine, 1979). This support may use either unilateral or bilateral tutoring. When teaching APE to students with severe disabilities, the most common strategy is unilateral. The literature

shows PALS has the potential to improve student engagement in learning locomotor skills as well as increase the number of attempts and the rate of response (Breslin & Rudisill, 2013). PALS have been proposed as best practice in pedagogy texts for general education and physical education (e.g., Siedentop & Tannehill, 2000) and as a method of inclusion for students with disabilities (Block, Oberweiser, & Bain, 1995). Although the research base supporting PALS in general physical education is growing, few studies have focused on APE students with severe disabilities in self-contained physical education settings (Ward & Ayvazo, 2006).

Peer tutoring with the use of typically developing peer tutors is emerging as an option to assist students with disabilities in the physical education setting (Block, 2007; Hodge, Lieberman, & Murata, 2012; Lieberman & Houston-Wilson, 2009). In this study, one typically developing peer takes on the role of the tutor, and the student with a disability is the tutee (d'Arripe-Longueville, Gernigon, Huet, Cadopi, & Winnykamen, 2002). Skilled peer tutors have been shown to yield better results in tutee skill acquisition than less skilled peer tutors (d'Arripe-Longueville et al., 2002; Johnson & Ward, 2001). Therefore, peer tutor training is critical for peer tutors and tutees to be as successful as possible in APE.

Although several studies have explored this topic in inclusive physical education settings, very few have focused on the self-contained setting (e.g., students with similar abilities are grouped in a classroom). The self-contained model necessitates peer tutors to be reverse mainstreamed (i.e., brought in from other classes) in order to have appropriate models for behavior, skill demonstration, etc.

The purpose of this study was threefold. The first purpose was to determine the

effects of peer tutor training on peer tutor teaching performance as measured by the Peer Tutor Evaluation Instrument. The second purpose was to determine the effect of peer tutors' support on tutees' improvement of motor performance, number of attempts to perform a motor skill, and number of steps taken (tutor and tutee) in a self-contained APE class. Third, this study sought to determine the social validity of the intervention based on practitioner feedback.

Literature Review

Peer Assisted Learning Strategies (PALS) include multiple instructional components that are supported by well-established learning theories (e.g., Piaget, 1985; Skinner, 1974; Vygotsky, 1978). PALS typically include peer modeling, collaborative activities, social interactions, assistance with motor skills, and positive and corrective feedback to support accurate skill acquisition. The following is a discussion of each of these characteristics with a connection to root learning theories from Piaget, Vygotsky and B.F. Skinner. Additionally, contact theory is discussed as it relates to attitude change for peer tutors in the SPE setting.

Peer Modeling

Peer modeling provides an opportunity for tutor and tutee to learn from one another. Ideally, the tutor demonstrates the skill correctly, and the tutee observes and executes the task that was just demonstrated. The tutor and tutee's skillsets are continually reconstructed. Piaget's learning theory of Equilibration is in line with how tutor and tutee interact while peer modeling is taking place.

Constructivism is a central theme in Piaget's Equilibration theory where

knowledge is built through continual reestablishment of what is known (Piaget, 1985). When students encounter events similar to what they know, the resulting information is assimilated into their existing knowledge. When students experience unique events, they may alter their thinking to accommodate the new experiences (Piaget, 1985). These two processes (i.e., assimilation and accommodation) often produce differing views of the world. In order for the student to reconcile these differences, they must change their position. Piaget called this self-regulation process "equilibration" (Piaget, 1985). Teachers who are willing to allow students to build knowledge through experience and collaborative activities support the Piagetian perspective of learning (De Lisi, 2002). Finally, Piaget suggests "interaction between tutor and tutee allows both reception and dissemination of information from both parties" (Piaget, 1985).

Collaborative Activities

Lave (1988) suggests that students acquire knowledge best through a culture of practice in realistic settings while working with peers. Rather than have students with disabilities repeatedly toss a bean-bag into a hoop, a potentially more successful situation would be to play a game where students decide which hoop to toss the bag into with reinforcement from a peer tutor (Ward & Ayvazo, 2006) in a collaborative manner. Moreover, peer tutors can help tutees confront learning at the "Zone of Proximal Development (ZPD)." The ZPD "is the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with *more capable peers"* (Vygotsky, 1978, p. 86). In essence, the playing of the game puts tutees in the ZPD. A peer tutor creates a realistic setting in which students learn the motor skill

as well as learn how to make decisions and interact with a peer in a collaborative setting.

Social Interactions

Vygotsky developed the cultural-historical theory that parallels Piaget's theory of learning in many ways. Vygotsky (1978) suggests that learning focuses on two primary characteristics, biological development and cultural mediation. Biological development refers to the regular maturation process that individuals experience, and cultural mediation is the conversion of social interactions to higher order thinking (Hogan & Tudge, 1999). Lave (1988) and Rogoff (1990), posit that cultural mediation (e.g., appropriate modeling) is embedded in many activities and social interactions students engage in when they are in pairs or small groups.

In many cases, the tutor may be able to offer feedback that is more age appropriate and accepted more readily than the practitioner. The social interactions between tutor and tutee may be more meaningful because they are more prevalent. These interactions also occur in a setting (e.g., one on one) that is more conducive to higher prevalence of interaction, leading to the natural maturation process (Vygotsky, 1978) of individuals.

Assistance With Motor Skills

Motor skill acquisition is often most successful for students when they are able to practice the skill correctly a high number of times. One method that ensures students are practicing correctly is immediate feedback. Students with disabilities should have immediate feedback regarding their performance of a task (Dunn & Leitschuh, 2010). B.F. Skinner offers support for immediate feedback from a behavioral perspective.

For example, peer tutors offer immediate feedback for attempts to throw a beanbag through a hoop. In this case, the student will begin to learn the correct technique through practice with positive and corrective feedback and a high number of attempts. The behaviorist perspective asks questions of the environment, in essence, evaluating current and historical markers (Skinner, 1974). Considering these two questions leads the practitioner to create situations where learning is most likely to occur. Feedback is another important aspect of PALS.

Feedback

Feedback from peer tutors creates opportunities for tutees to begin to discriminate and generalize, two skills important for social development. Discrimination and generalization are learning processes described by Skinner. Discrimination is to behave differently in similar situations that have slightly different characteristics, such as kicking a ball with the instep in a practice session, but then kicking the ball with the toes in a game-like situation. A tutor can model discrimination by kicking the ball with minimum force to a peer who is 5 feet away and with a lot of force to a peer who is 30 feet away.

To generalize is to behave similarly in different situations. For example, a student might demonstrate a mature pattern for kicking a ball in the gymnasium and then demonstrate the same mature pattern for kicking a ball on the soccer field. This student is demonstrating generalization by exhibiting the same skills in two different environments. Peer tutors can help tutees generalize skills through the use of differential feedback.

Differential feedback is provided when a tutor gives positive feedback for correct

performance and corrective or no feedback for incorrect performance (Ward & Lee, 2005). If a tutor's differential feedback (e.g., praising the correct performance of kicking the ball with the instep to pass, but not the incorrect performance of kicking the ball with the toes) works to improve tutee performance, then the tutor becomes an integral part of the learning environment and process.

Contact Theory

Particular emphasis on ensuring meaningful interaction between parties emerges from examining contact theory. Moreover, casual contacts (e.g., a student with severe disabilities disrupting a classroom while running down the hall) are more likely to increase prejudice than dispel it (Allport, 1954). Hence, every casual contact made with an outgroup member (student with disability) theoretically strengthens the adverse mental associations held because people see what they want to see (i.e., what people have learned to see) (Slininger, 2000).

Contact and acquaintance programs (e.g., peer tutor programs) have been suggested as a method of reducing stereotypes, prejudice, and discrimination (Slininger, 2000). However, they must meet criteria, (1) instill a sense of equality and social status, (2) embody ordinary pursuits (e.g., skills to be active with friends), (3) be set in a realistic manner, (4) and have the support of the greater community (e.g., recess buddies, other instances of academic peer tutoring for students with disabilities, etc.) (Allport, 1954).

Peer tutor programs should be designed to correct stereotypes and facilitate positive beliefs. Once these criteria are met, the trend favors the idea that knowledge about minority groups make for tolerant and friendly attitudes (Slininger, 2000). One cannot assume the relationship is flawless; and it cannot be clear whether knowledge

causes friendliness; however, there are positive relationships between knowledge and friendliness (Allport, 1954).

Contact theory was conceptualized as a broad approach to affecting intergroup relations (Slininger, 2000). This broad approach follows the idea that knowledge about minority groups must be built through direct experience and credible information sources. Consequently, there are five categories of variables to be carefully considered when designing research concerning the effect of contact on attitudes; (1) quantitative aspects of contact (i.e., frequency, duration, number of participants), (2) aspects of contact, (3) social atmosphere surrounding contact, (4) personality of the participant, (5) areas of contact (Allport, 1954).

Research studies are conducted in an attempt to provide answers to how learning occurs and what likely has contributed to the learning (Ward & Lee, 2005). Consideration of root theory from Vygotsky, Piaget, Skinner, and Allport provides perspective from which to view current learning practice, especially the use of peer tutors in adapted physical education. The characteristics of PALS, (e.g., peer modeling, collaborative activities, social interactions, assistance with motor skill development, and feedback) are rooted in the learning theories from Vygotsky, Piaget, Skinner, and Allport. The use of these theories to guide the study provides an important framework for research and understanding how students learn with peer tutors.

Peer Tutors

Students have the opportunity to serve as instructional agents for one another through peer mediated instruction (Harper et al., 1994). Peer tutoring is a potential approach for addressing essential needs of students in the special education classroom

(McDonnell, Thorson, & Allen, 2000). Peer-mediated instructional strategies allow students to learn from one another, even though one student is the tutor and the other is the tutee (McDonnell et al., 2000).

Research on peer tutoring is extensive in both general and special education classroom settings (Mcdonnell et al., 2000). The effects of peer-mediated instruction in the classroom for students with severe disabilities have been positive (McDonnell et al., 2000). Literature also show that students without disabilities can be successful peer tutors for students with severe disabilities in the classroom (Carr, 2005; Fenrick & McDonnell, 1980; Kamps, Locke, Delquadri, & Hall, 1989), and the special physical education classroom (Lieberman, Dunn, van der Mars, & McCubbin, 2000; Webster, 1987; Wiskochil, Lieberman, Houston-Wilson, & Peterson, 2007). Moreover, students without disabilities appear to gain social benefits from the peer tutoring programs, including empathy, improved self-esteem, and friendships with students with disabilities (Owen-DeSchryver et al., 2008). In addition, the peer tutoring programs often illicit increases in the total duration of social interaction between groups of students during free time (e.g., students saying hello in the hall, or students eating together in the lunchroom) (Fenrick & McDonnell, 1980).

There has been less research addressing attitude change for peer tutors participating in a peer tutor training program and working with students with severe disabilities in adapted physical education (Slininger, 2000). However, attitudes of peers can be one of the most important variables in the success of a peer tutor dyad (Lieberman et al., 2000). Not all peer tutors have excellent attitudes toward their peers with disabilities. However, training for peer tutors can marginalize some of these negative

attitudes (Klavina, 2008).

Attitude is a complex construct with several definitions (Ajzen & Fishbein, 1980; Antonak, & Livneh, 1988; Tripp & Sherrill, 1991). The construct of attitude has been restricted to an individual's evaluation of a psychological item (Ajzen & Fishbein, 1980). Other theorists consider attitude as being multidimensional and focus on the importance of measuring cognitive, affective, and behavioral aspects. Triandis describes an attitude "as an idea (cognitive component) charged with emotion (affective component) which predisposes (cognitive component) a class of actions (behavioral component) to a particular class of social situations" (1971, p. 2). Sherrill (1998) streamlines the concepts of Triandis (1971) by describing attitude as: "… an enduring set of beliefs charged with emotion that predisposes a person to certain kinds of behaviors" (p. 7).

Although definitions may vary, there seems to be agreement in the literature that attitude research should be based on a specific set of six characteristics. The first characteristic is that attitudes are learned through experience and interaction with other people, (2) attitudes are complex, (3) attitudes are resistant to change and relatively stable, (4) attitudes have a specific social object as a referent, (5) vary in their quantity and quality, possessing different degrees of motivating force (intensity, strength), and direction (toward, against, away from the attitude referent), and (6) are manifested behaviorally via predisposition to act in a specific way when the individual encounters the attitude referent. These six characteristics provided the framework for this study of peer tutor attitudes in SPE toward their tutees with severe cognitive disabilities (e.g., the specific object of referent).

Slininger et al. (2000) suggest contact theory as a guiding principal for studying

attitudes in SPE. This study assumes that the students who are significantly different from the normative group (e.g., students with severe disabilities) may be marginalized in the educational setting (Dunn, & Leitschuh, 2000; Henderson, Lavay, & French, 2006; Slininger et al., 2000). Marginalization can result from various degrees of negative behavior, avoidance, discrimination, physical attack, etc.

Contact theory provides a framework from which to study the relationships between tutor and tutee. However, groundwork should be set regarding the theory. Contact theory suggests that prejudice and discrimination toward a minority group will be reduced when contact between individuals meets four criteria: (1) equal status among parties, (2) community support for the change, (3) common objectives should be sought by both parties, (4) the connection must be genuine and trusting (i.e., tutors should know their tutee well, and be interested in their success in the setting) (Allport, 1954).

Particular emphasis on ensuring meaningful interaction between parties emerges from examining contact theory. Moreover, casual contacts (e.g., a student with severe disabilities disrupting a classroom while running down the hall) are more likely to increase prejudice than dispel it (Allport, 1954). Hence, every casual contact made with an outgroup member (student with disability) theoretically strengthens the adverse mental associations held because people see what they want to see (i.e., what people have learned to see) (Slininger, 2000).

Peer tutor programs should be designed to correct stereotypes and facilitate positive beliefs. Once these criteria are met, the trend favors the idea that knowledge about minority groups makes for tolerant and friendly attitudes (Slininger, 2000). One cannot assume the relationship is flawless and it cannot be clear whether knowledge

causes friendliness; however, there are positive relationships between knowledge and friendliness (Allport, 1954).

This study examines a partner learning process in which one student with a disability and one typically developing peer work together to gain selected locomotor and object control skills and increases in physical activity. The process is designed to ensure that students with disabilities receive instruction tailored to their individualized educational needs within the typical structure of the APE setting, and to maximize the interactions with typically developing peers. The study evaluated partner learning in terms of effects on tutee motor performance (e.g., criteria met for a specific skill), tutor feedback (e.g., number and type of feedback provided by tutor), and tutor and tutee physical activity (e.g., step counts).

Statement of Problem

This study extended the research base for peer assisted learning strategies (PALS) as an approach for supporting students with severe disabilities in APE. There is significant research surrounding PALS in general (Charlop, Schreibman, & Tyron, 1983; Fenrick & McDonnell, 1980; Kamps, Locke, Delquadri, & Hall, 1989; Lancioni 1982), and special education settings, (Greenwood et al., 1991) as well as inclusionary physical education settings (Lieberman, Dunn, van der Mars, & McCubbin, 2000; Owen-DeSchryver, Carr, Cale, & Blakeley-Smith, 2008; Webster, 1987; Wiskochil, Lieberman, Houston-Wilson, & Peterson, 2007). However, this study will be one of two (Breslin & Rudisill, 2013) focusing on students with severe disabilities in self-contained settings. Currently, there is not a peer tutor evaluation instrument and training protocol for teacher use to evaluate and train peer tutors in APE.

Purpose Statement

The purpose of this study was threefold. The first purpose was to determine the effects of peer tutor training on peer tutor performance as measured by the Peer Tutor Evaluation Instrument. The second purpose was to determine the effect of peer tutors' support on tutees' improvement of motor performance, number of attempts to perform a motor skill, and number of steps taken (tutor and tutee) in a self-contained APE class. Third, this study sought to determine the social validity of the intervention based on practitioner feedback.

Research Questions

1. Does training peer tutors positively affect their teaching behaviors in Adapted Physical Education (e.g., engagement and interaction, use of peer tutee name, prompting, feedback, and social interaction)?
2. Do students with severe disabilities taught by trained peer tutors show improvement in motor skill development, and/or have higher levels of activity when compared to a condition in which the students with severe disabilities worked with the same peer tutors prior to training?

Method

Participants

Twelve participants with disabilities, 7 boys and 5 girls age 6-12 (M = 9.08, SD = 2.11), were randomly selected from four self-contained special education classrooms to take part in this study. Each of these participants had a current Individualized Education Plan (IEP) at the time of the study; disabilities included intellectual disability, ASD, other

health impairment, multiple disabilities, and/or developmental delay. All participants with disabilities are enrolled in a self-contained adapted physical education class for 30 minutes per week. Four classes participated in the study.

Twelve typically developing students from the same school were randomly selected from four classes (3rd and 6th grades) to participate in the study (5 boys and 7 girls) age 9-12 (*M* 10.33, *SD* 1.44) recruited to be peer tutors for this study. Peer tutors were released from academic learning sessions, hence classroom teachers selected students to participate based on merit (e.g., academic performance, citizenship) and willingness to participate. If at any time the classroom teacher noticed a negative effect on peer tutor academic performance, they were released from the peer tutor program – no peer tutors were released. Peer tutors were assigned to dyads based on special education teacher suggestions. These dyads were static throughout the course of the study.

Materials and Procedures

The *Test of Gross Motor Development* – 2 (Ulrich, 2000) was used to evaluate tutee performance on selected motor skills for students with disabilities. This test is norm referenced and used widely for assessing motor skills of children with and without disabilities ages 3 to 10. Students with disabilities participated in a pretest and posttest assessment. Pretest occurred during the second session of the study. Posttest was administered during the 14th session. Both assessments were video recorded and scored following the session.

The number of steps for both peer tutors and tutees were measured using CW-701 Yamax Digi-Walker pedometers. This model is a small plastic box that was clipped to the waistband. The pedometers provided an objective assessment of total steps taken in a

set amount of time for each individual (Welk, 2002). Pedometers were placed on the peer tutor and tutee at as close to the same time as possible. Steps were recorded for the entire study session.

Peer tutors performance was evaluated using the Peer Tutor Evaluation Instrument. This instrument consists of 7 goals and 9 objectives used to evaluate peer tutor performance during the tutoring session. This instrument was developed by the principal investigator and deemed to have content validity by a panel of expert reviewers. The goals and objectives in the instrument focus on teaching qualities for peer tutors.

Peer tutor attitude was estimated using a smiley face scale (Kunin, 1955). The "smiley" scale has been most commonly used with children in research environments. Wells (1965) presented the "smiling face" scale as a liking measure for children 5-12 years old because they are more apt to identify with faces than words or numbers (Macklin, & Machleit, 1990). Practitioners report the use of the smiley face scale as a viable option for efficient assessment of attitude toward a specific construct (Kunin, 1955). The smiley face scale takes very little time to introduce because it is ubiquitous with elementary school settings (Sad, 2012). Given the time constraints, it was decided to pair the smiley face scale with a verbal response to a specific prompt at the end of each teaching session to estimate peer tutor attitude.

Social validity was evaluated using the Intervention Rating Profile – 15 Modified (IRP-15-Modified) (Martens & Elliot, 1985). This rating profile was used to provide general and special education teachers cooperating with the study an opportunity to rate the validity of the intervention. The IRP-15 is an instrument created to be modified by the researcher (Witt & Elliot, 1985). Two versions of the IRP-15 were created for this

study. General education teachers (4) who cooperated with the study by recruiting and providing peer tutors were given the "IRP-15-Gen Ed." Special education teachers (5) whose students were the tutees for the study were given the "IRP-15-SPED." Both versions focused on the peer tutor training program; however, each was modified to access the unique perspectives of general and special education.

One Apple iPad video camera was used to video record each data collection session. A conspicuously positioned graduate assistant in the adapted physical education setting operated the iPad. The decision to use the iPad was made after consideration for the need to move around in the space in order to focus on each work station, or a specific tutor dyad. iPads are also ubiquitous in the special education classrooms and it was determined that its presence in the space would be less distracting than a video camera.

Wireless voice recorders (Sony ICD-PX440, Sony ICD-UX533) fitted with lavaliere microphones were used to record verbal interaction between tutor and tutee. The voice recorder was placed in the peer tutor's pocket. The microphone recorded audio within close proximity (5 feet) of the tutor. These devices were used to provide clear audio data of peer-to-peer interaction in a loud physical activity setting.

Number of steps for both peer tutors and tutees were measured using CW-701 Yamax Digi-Walker pedometers. This model is a small plastic box that was clipped to the waistband. Pedometers provide an objective assessment of total steps taken in a set amount of time for each individual.

Procedure

Data collection occurred in the SPE teaching site. Participants completed assent forms prior to the beginning of the study and were reminded that they may abandon the

study at any time. Three randomly selected peer tutors (drawing names out of a container) had the voice recorder in their pocket and wore the lavaliere microphone clipped to their shirt during each study session – these selections remained the same throughout the course of the study. Both tutor and tutee wore a pedometer attached to their waist with a plastic clip. Pedometer data for tutor and tutee were recorded after each teaching session. Audio data were coded from recordings after the sessions were completed for the day. Attitude data were collected on the "smiley" form after the completion of the class. Students stated the class, date, and an assigned letter (designation for their name) into the audio recorder before the data collection session began for identification purposes.

All study sessions took place in a multipurpose room, approximately 50 feet long and 30 feet wide. The room was the same one in which the students currently participate in SPE. There was a schedule of activities for the day posted conspicuously on a wall for reference by the tutors, staff, and the teacher throughout the study.

The peer tutors participated in one 30-minute training session (intervention), and 2-5-minute booster sessions prior to each class they tutored in. These sessions focused on learning and honing skills to be an effective peer tutor. The Klavina et al. (2008) training protocol has been adapted as a model for training peer tutors in this study. These training sessions are designed to offer tutors instructional assistance to the peer with severe disabilities in a clear and understandable manner. Peer tutors also participated in a 2-4-minute booster session focusing on teaching objectives at the beginning of each class. Peer tutor objectives were conspicuously placed near teaching stations for reference during the session once intervention had been implemented.

Design and Analyses

The data were collected using a single subject multiple baseline design (Shadish, Cook, & Campbell, 2002).

Statistical Analyses

Paired samples *t*-test (Field, 2013) was used to evaluate pretest and posttest scores in order to determine if there was a statistically significant difference in pretest and posttest scores for all groups for TGMD-2 skill performance.

Peer tutor attitude ratings were analyzed using the Wilcoxson Signed Ranks (Field, 2013) test to determine if there was a statistically significant difference in negative and positive comments when baseline and intervention were compared.

Motor Skill Assessment

The motor skills that were assessed in this study were determined by a pretest assessment using the *Test of Gross Motor Development 2* (TGMD-2). The TGMD-2 (Ulrich, 2000) was designed to assess the motor skill performance of students ages 3-10. The assessment for each skill is broken into varying numbers of criteria (3-4) as described by the TGMD-2.

Each motor skill was assessed for the presence/absence of each of the criteria defined in the TGMD-2 as a mature (e.g., correct execution of the skill as defined by TGMD-2) pattern for that skill. The sum of all trials during each assessment was plotted to show the level of mastery for each skill for each participant. In order to make the assessment periods as equal as possible, each participant will have two trials/attempts to complete the skill during each assessment period.

Two locomotor (skip, gallop) and two object control (underhand roll, overhand throw) skills were then selected based on assessment outcomes, potential for improvement, and generalizability of that skill to other settings. All participants in the study were taught the same four skills.

Peer Tutor Performance

Research (Klavina, 2000) shows that feedback from peer tutors can have a positive impact on skill acquisition for peer tutees. Peer tutors were trained in specific methods of teaching motor skills, including communication, prompting, and providing feedback. Untrained peer tutor performance was evaluated in the baseline phase. Once they were trained (intervention), they were assessed on their teaching performance during the skill session using the same instrument.

Pedometer Data

Tutors and tutees wore pedometers for the duration of each session (16) in the study. Pedometers were marked to ensure tutor and tutee step counts can be paired. Pedometers were checked with a shake test (Welk, 2002) prior to each study session, and were placed on tutor and tutee as close to the same time as possible. Tutor and tutee step count data were recorded at the end of each class and were plotted. Tutee step data were used as a comparison measure to evaluate any correlation between tutor performance and tutee activity during data analysis.

Each study session was video recorded in its entirety, although only the practice portion of the class was used for peer tutor performance analysis. The practice session consisted of instruction in the four selected motor skills at two different stations.

Graduate students taught preset *I Can Primary Skills* (Wessel, & Zittle, 1998) curriculum at each station.

Observer Reliability

In order to determine interobserver agreement for the Peer Tutor Evaluation Instrument, 30% of the class sessions (5) were randomly selected to be reviewed by two independent observers (i.e., the author and research assistants). Independent observers' scores were compared with those of the author on an item by item basis. Both observers would have to score each criterion the same way (i.e., if the peer tutor states the tutee name at least four times in the course of the practice session, both observers would have to agree on the number of times the tutor said the tutee name and score a four on that objective). Interobserver reliability was calculated as the number of agreements/agreements plus disagreements, multiplied by 100. A predetermined mean of ≥ 80% interobserver reliability was deemed acceptable. Analysis included reviewing video and audio recordings to complete the Peer Tutor Evaluation Instrument (PTEI) for each peer tutor participating in the class. Observers were allowed to review recordings as many times as necessary to complete the PTEI. Once the PTEI was completed, scores were compared and reported IOA was calculated. The range of interobserver reliability was between 78% and 100% with a mean of 85.6%.

TGMD-2

When measuring interobserver agreement (IOA) for the TGMD-2 pretest and posttest assessments, a trained independent observer needed to concur that the same number and type of TGMD-2 criteria were present in each trial for an agreement to be

scored. For example, if the participant completed a trial of the underhand roll and performed all of the criteria except for following through, they would score a three out of a possible four for that trial. Both observers would have to agree that the same criteria were achieved or missing from the trial. Agreement was calculated as the number of agreements/agreements plus disagreements, multiplied by 100. A predetermined mean of ≥ 80% interobserver agreement was deemed acceptable. IOA was calculated for both the pretest and posttest assessments. Video recordings for both pretest and posttest phases of the TGMD-2 were reviewed by a trained observer. The observer reviewed the footage and scored the number of criteria present for the skill. Scores were then compared with those of the author and IOA was calculated. The range of IOA was 94% to 100% with a mean of 96.7%.

Peer Tutor Attitude

When measuring interobserver agreement (IOA) for the qualitative response data, a trained independent observer needed to score each response as a positive or negative. Each entry in the spreadsheet was deemed to be a positive or a negative comment by both the graduate assistant and primary researcher. Once the data were audited agreements and disagreements were tallied. Agreement was calculated as the number of agreements/ agreements plus disagreements, multiplied by 100. A predetermined mean of ≥ 80% interobserver agreement was deemed acceptable. IOA was 86%.

CHAPTER 2

STUDY 1: DEVELOPMENT AND INITIAL VALIDATION OF THE PEER TUTOR EVALUATION INSTRUMENT

Abstract

The purpose of this study was to delineate the development of a peer tutor evaluation instrument to assess interaction and teaching skills of peer tutors (5 males, 7 females) age 9 to 12 (M = 10.3, SD = 1.4) in adapted physical education. The theoretical and empirical base for the evaluation instrument is detailed in a three-step study. The first step was to develop 7 goals and 9 objectives, based on literature review, consultation with professionals in the field, and personal experience. The second step involved engaging expert reviewers to determine content-related evidence of validity (CV) of the goals and objectives on the evaluation instrument as they relate to the peer tutor training program. The third step involved using the evaluation instrument in practical settings to assess reliability and usability. The outcomes of content validity and reliability ratings favorably indicated that the Peer Tutor Evaluation Instrument (PTEI) is an effective tool for practitioners to use for rating peer tutor performance in adapted physical education settings.

Introduction

Peer tutoring is a commonly utilized resource when there is a need to provide more support for a student in educational settings. Often, the peer tutor provides additional support and practice for a student who may be underperforming. These same scenarios can be found in diverse settings such as extracurricular activities, competitive sports, and physical education classes. In these settings, students should be gaining essential skills for life-long activity; however, students often struggle learning motor skills due to a lack of resources (inadequate staff, limited equipment) and large class sizes. In order for all students, including those with disabilities, to be successful,

practitioners need additional assistance to teach these essential skills and have students practice them. Peer tutoring is a common method of providing that assistance to improve student learning outcomes in physical education (Block, 2007; Hodge, Lieberman, & Murata, 2012; Kamps et al., 1994).

Peer-assisted learning strategies (PALS) is a program conceived by Rosenshine (1979) in which peer tutors are used to support direct instruction. Instruction is presented by peer tutors who typically have a higher skill level and aptitude for working with other students (Block, 2007). Typically peer tutors will be trained; however, the peer tutoring can be formal (e.g., students using a curriculum book and teaching lessons) or informal (e.g., students reading to one another, working on math problems, etc.). PALS are usually administered utilizing one of two common strategies. Bilateral tutoring occurs when students learn content from one another. Peer tutor dyads in this situation are usually similar age or ability. Unilateral tutoring is a scenario where one student is responsible for teaching content to another student with a lower level of skill (Fenrick & McDonnell, 1980).

PALS have been proposed as a best practice in pedagogy texts for general education and physical education (e.g., Siedentop & Tannehill, 2000) and as a method of inclusion for students with disabilities in physical education (Block, Oberweiser, & Bain, 1995). Ensergieux and Lafont (2010) used the PALS program in a unilateral design to affect motor performance and self-efficacy outcomes for students with severe disabilities. They found unilateral peer tutoring positively influenced motor performance and self-efficacy for tutees.

Peer Tutor Evaluation

A major caveat to the use of peer tutors in physical education is that they should be trained and evaluated to ensure their competence in teaching the assigned tasks (Houston-Wilson, van der Mars, & McCubbin, 1997; Lieberman, van der Mars, & McCubbin, 2000; Wiskochil, Lieberman, Houston-Wilson, & Petersen, 2007). There are several informal peer tutor evaluation tools for inclusive physical education (Lieberman, & Houston-Wilson 2012); however, these assessments tend to focus more on Academic Learning Time – Physical Education (ALT-PE) in the session (Webster, 1987) rather than peer tutor performance. ALT-PE is a systematic evaluation with specific categories (motor appropriate, motor inappropriate, motor supporting, not motor engaged) evaluating activity levels. These categories are systematically observed and students are evaluated on whether or not they exhibit the behaviors. Frequency scores are tallied for the session.

Another instrument, Analysis of Inclusion Practices in Physical Education form S (AIPE-S; Hodge et al., 2000), was designed to evaluate occurrence and duration of specific student interaction for eight student behaviors. The behaviors include initiation of communication, modeling tasks, praise for effort, name use, appropriate feedback, provides "hands-on" help, other interactions, and no interaction. The AIPE-S differs from the peer tutor evaluation instrument discussed here because it has objectives written to evaluate both peer tutor and tutee.

While these tools are effective in evaluating tutee performance or tutor/tutee interactions, there was no validated evaluation instrument to objectively measure peer tutor interaction and tutoring performance. This new instrument could also serve to

identify specific areas that need to be included in the training and preparation of peer tutors in order for them to be more successful tutoring children with disabilities in physical education.

Peer Tutor Training

Realizing the importance of training peer tutors, Klavina (2008) developed a peer tutor training protocol. This protocol was revised to include language and examples focusing on elementary age students, and to include booster sessions. The rationale for revising the training was to engage peer tutors as much as possible, and to have booster session topics to use before APE class to help peer tutors focus on improving their tutoring skills. The goals and objectives of this instrument were matched to the new peer tutor training protocol, in order to ensure tutors were trained on the items on which they would be evaluated.

The Klavina (2008) training protocol was initially designed with three individual training sessions of 30 minutes. We amended the Klavina training to include one 30-minute training session plus brief booster sessions before each tutoring session. This change in protocol was required due to significant time constraints on peer tutors (e.g., academics, testing, specialty classes). The format of this training was a favorable solution to the time constraints placed on the peer tutors. Moreover, it may be more successful with younger students due to limited attention spans.

The Peer Tutor Evaluation Instrument (PTEI) is an observation instrument that was specifically designed to obtain data on the tutoring and interaction performance of peer tutors. The purpose of this article is to discuss the development, assessment of validity evidence, and preliminary use of the PTEI. The research question this study

sought to answer was the following: to what degree is the PTEI rated as having adequate content validity and reliability by experts in APE? The hypothesis for the study was the new instrument would be rated as having high content validity and reliability.

Method

This study followed a two-stage process to create the PTEI and establish content validity and reliability for it.

Stage 1: Development of the PTEI

Lynn's (1986) instrument development stages were used to develop the PTEI. Stage 1 focused on developing the instrument. This development stage included the identification of content domain, element creation, and instrument formation (Lynn, 1986). Once the development stage was complete, evaluation and quantification of the content validity of the instrument through expert review began (Stage 2). The steps that were followed for creating the instrument and evaluating content validity (CV) are delineated in Table 2.1.

Additionally, the authors conducted an evaluation of interrater reliability, asking research assistants to use the instrument and compile results. Initially, the research assistants were trained to use the instrument by viewing practice video recordings. The same research assistants then reviewed video and audio recordings of peer tutors from practice sessions and scored the performance of specific peer tutors. Scores were compared once both assistants had completed rating the peer tutors. Interrater reliability was calculated for item agreement.

The foundation of the PTEI originated with a literature review concentrating on

peer tutor training and evaluation protocols. Databases searched included EBSCO, SCOPUS, and SportDISCUS. Colleagues in the field were contacted to ascertain if specific observation and evaluation instruments were used in their peer tutor programs. Additionally, the researcher has 5 years of experience teaching SPE. During this time, the researcher collected anecdotal notations of interaction behaviors between students with and without disabilities in instructional settings. Theoretical and empirical information in the literature and personal experience were used to inform the preliminary list of PTEI contents.

Selection and Element Creation

Goals and objectives on the evaluation instrument were selected based on the identification of recurring categories in the literature (Lieberman, 2013), and practical experiences focused on identifying essential skills for peer tutors. The three following main categories were identified: 1) communication between peer tutor and tutee, 2) peer tutor prompting, and 3) peer tutor feedback (d'Arripe-Longueville et al., 2002; Houston-Wilson et al., 2009; Klavina, & Block, 2008; Lieberman, 2013). Categories were divided into seven goals (see Table 2.2). These goals relate directly to the main content categories (e.g., communication, prompting, feedback; Lynn, 1986). Nine measurable objectives directly correlating to the peer tutor training protocol were written. The instrument was then developed and refined. Instrument development involved formatting the instrument for clarity, usability, and efficiency. Refinement included integrating feedback from experts regarding utility, ease of use, and suggested edits were considered and applied to the PTEI.

Stage 2: Content Validity and Instrument Review

The categories in the PTEI are communication, prompting, and feedback that the peer tutors use while they are working with their tutee. Each category is further broken into goals and objectives to garner a clearer image of peer tutor performance in the APE setting (see Table 2.2). The PTEI instrument (Figure, 2.1) shows how goals and objectives were broken down and added to the instrument. The researcher considered the limitations (e.g., time, large class sizes, limited support, etc.) facing practitioners in the field when developing the PTEI. Hence, the instrument includes a limited number of goals and objectives. These measures serve to quantify peer tutor actions toward their tutee.

Communication behaviors are characterized by how well the peer tutor communicates with the tutee once they are paired in the SPE setting. These behaviors include proximity, asking check-in questions, and using the tutee's name multiple times during the session. Although we considered other modes of communication (e.g., nonverbal, etc.), we decided that these three objectives focused on the main modes of communication to be evaluated by the peer tutor instrument.

Motivating students with disabilities is an essential aspect of tutoring in APE. Tutors need to have the ability to correctly prompt their tutee to perform a task (Klavina, 2008). Prompting behaviors are any verbal instructions that the peer tutor directed to the tutee related to performing a skill or task, waiting at least 10 seconds for the tutee to respond, and using physical prompts when the tutee does not respond (Klavina, 2008; Ensergueix & Lafont, 2010). A physical prompt in the APE setting differs from the special education setting. In APE, it is often helpful for the tutee to be prompted with a

"hand-over-hand" prompt so they feel the correct motor movement associated with throwing a ball overhand, kicking the ball with the instep of their foot, spreading their fingers to catch a ball, etc. (Dunn & Leitschuh, 2010). Although a physical prompt in the special education setting may address noncompliance (e.g., a student refusing to participate), aggressive behavior (e.g., hitting, kicking, biting, etc.), or safety issues (e.g., keepings students safe from one another, off tables, desks, throwing objects), peer tutors should not be physically prompting tutees who are noncompliant or misbehaving (e.g., picking their tutee up off the floor because they are refusing to participate, etc.). In this case, the tutor should contact the teacher or staff to offer assistance (Lieberman, 2013).

Feedback behaviors include the use of different types of feedback to help the tutee learn and perform the task. These types of feedback may include positive specific performance feedback, positive specific behavioral feedback, positive general feedback, nonverbal feedback, and corrective feedback. In positive specific performance feedback, the tutee is told specifically how well the task was performed, such as, "Johnny, you kicked the ball with the inside of your foot." Positive specific behavioral feedback addresses the tutee's social behavior, such as, "Johnny, I really like how you listened to me and picked up the ball to throw it." With positive general feedback, the tutee is provided a general statement of approval for doing the task, such as "Wow! Nice work!" Nonverbal feedback is providing approval through manual gestures such as high fives or head nods. Finally, corrective feedback is telling the tutee they need to improve a skill (Klavina, 2008). Positive specific performance and positive specific behavioral feedback are the two types of feedback that were selected to be included on this instrument because they are important tutoring skills for tutors to learn (Lieberman, 2013).

Frequency of social interaction behaviors was included in the instrument. These behaviors included verbal or nonverbal communication on content not related to APE such as comments on dress, comments to engage, and compliments. Even though these behaviors are not directly related to tutee performance in APE, they may affect social acceptance and relationships between students with disabilities and their typically developing peers (Block & Malloy, 1998; Tripp et al., 1995).

Design

A Likert (Likert, 1932) scale survey design was used in this study to establish content validity by experts. This design establishes validity through the scores of expert reviewers agreeing or disagreeing with the appropriateness of the item and its relevance to the topic. Potential weaknesses in this design are a small number of participants as expert reviewers ($N = 5$). A low sample size requires a high rate of concordance between raters' scores.

Procedure

Content validity of a new instrument considers how well the conceptual domain is represented by specific pieces of the instrument (Cook & Beckman, 2006). Content validity considers logical rather than statistical features of the instrument (e.g., the wording of a goal or objective, organization of the instrument; Klavina, 2008). Establishment of content validity is typically presented as a detailed description of steps taken to make sure the items correctly represent the construct (Lynn, 1986). Typically, multiple experts rate each item in terms of its relevance to the content (Lynn, 1986).

The content validity of the PTEI was addressed by following a 3-step process: 1)

identification of content domain, 2) identification of content domain items, 3) evaluation of content and items based on specific criteria. Five internationally known experts (professors and assistant professors) in APE were contacted and asked if they would be willing to evaluate the degree to which the items in the PTEI related to a defined domain of content (Yun & Ulrich, 2002). Experts were from the United States and Europe with well-respected reputations in the area of APE and having publications focusing on behavior management, peer tutors, or evaluating student performance. Three of the selected panelists had experience developing behavior measures in the physical activity domain. These experts were selected to provide a significant depth and breadth of APE knowledge in the content validity and rating process of the PTEI. All experts who were contacted agreed to rate the PTEI. Changes suggested by the experts were then made to the original PTEI on the criteria of scoring, wording of objectives, and the organization of the objectives.

To address the content validity of the PTEI, the following steps were taken by the expert reviewers: 1) read intended use and scoring procedures for the instrument; 2) determine the degree to which each item on the peer tutor evaluation instrument (PTEI) is correlated to the corresponding learning objective on the peer tutor training protocol (PTTP); 3) complete the rating scale, along with adding any comments for editing format, etc.

Evaluation of Items

An evaluation scale for experts to complete was designed. This scale contained the goals and objectives in a table, as well as the PTEI. Experts were asked to rate the relevance or match to the specific goal and objective in the peer tutor training program

(PTTP) on the evaluation scale. The 4-point Likert scale (1 = no match at all, 4 = excellent match) was used to rate the relevance. It was expected that the goals and objectives would be rated as an excellent match because the Peer Tutor Evaluation Instrument and Peer Tutor Training Protocol share the same goals and objectives.

Expert reviewers were asked to rate each evaluation criterion on a 1-4 scale (1 = the evaluation objective does not correlate with PTTP objective at all; 2 = unable to assess without item revision; 3 = correlates but needs minor alteration; 4 = excellent correlation of items; Lynn, 1986). A 4-point rating scale is preferred because it eliminates the ambivalent middle rating (Lynn, 1986). Expert reviewers were also asked to identify any areas that had been omitted from the instrument.

Item Relevance

Experts who were in agreement about an item's relevance (scored an item as 3 or 4) presented a quantitative measure of content validity. Kazdin and colleagues (1982) suggested that an average item rating agreement of 80% is acceptable. A percentage agreement was calculated based on expert ratings. Percentage agreement was interpreted as follows; 20% to 40% was considered low concordance, 41% to 60% as moderate concordance, 61% to 80% as good concordance, and 81% to 100% as excellent concordance (Kendall & Gibbsons, 1990; Siegel & Castellan, 1998). This percentage was calculated for agreement and relevance for each item.

Research assistants were asked to evaluate the PTEI in a similar manner. They were also asked to use the PTEI in an authentic setting and offer feedback on the utility of the instrument and evaluate whether or not the items represent the domains of communication, prompting, and feedback. A 5-item questionnaire was designed with a

4-point Likert scale (Likert, 1932) that was used to evaluate user opinion and feedback.

Results

Stage 2: Content Validity and Instrument Evaluation

Feedback from expert reviewers resulted in changes in scoring procedures, and creating a percentage of performance outcomes for ease of interpreting the results. Initially, the instrument used a mix of frequency (i.e., how many times the peer tutor used the tutee's name), and qualitative measures (i.e., how well did the tutor provide performance feedback). One expert suggested that the scoring be revised to eliminate the qualitative measures and use only frequency scoring for ease of use. Another expert suggested using a percentage of total possible points to summarize the scores, rather than just a total score.

The results of stage 2 are presented in Table 2.3. The magnitude of overall percentage agreement was (M = 3.88) out of 4, which indicated high agreement on ratings of content validity from expert reviewers.

Feedback from research assistants (n= 2) concurred with experts' reviews (see Table 2.3), with a similar percentage agreement of 97%. Qualitative feedback from research assistants (reported in Table 2.3) was also considered in the final iteration of the PTEI.

Discussion

This study involved the development and initial content validity outcomes for the PTEI, an evaluation instrument designed to measure peer tutor teaching and interaction performance in adapted physical education settings. The PTEI items were established

using existing literature in both adapted and general physical education settings as well as professional experience. Ratings by expert reviewers and graduate assistants provided essential feedback for the development and refinement of the PTEI.

Expert reviewers provided content-related evidence of validity for the PTEI. A majority of the expert reviewers rated each PTEI item as relevant and concurring with the defined categories of communication, prompting, and feedback. All results showed high levels of agreement (Table 2.3). The lowest mean and percentage of agreement across evaluators was for prompting. This lower percentage of agreement suggests that future iterations of the PTEI may need additional evaluation of the prompting goals and objectives. Additional evaluation of these goals and objectives may create a more robust instrument.

The PTEI instrument has limitations. It contains only nine objectives for evaluating the myriad behaviors and interactions that occur between tutor and tutee in the APE setting. This instrument is designed to provide an evaluation of peer tutor performance, but does not encompass the entirety of variables present in the situation. The goals and objectives selected for the PTEI were determined by literature review, previous peer tutor training protocols (Klavina, 2008) and practitioner handbooks (Houston-Wilson et.al., 1997; Lieberman, Houston-Wilson, 2009). These resources presented a wealth of information, which was distilled into the PTEI by writing goals and objectives that best reflected the major themes in the review of these resources.

Conclusions

The results of this study provide preliminary evidence of content validity, in elementary APE. The PTEI offers an important assessment tool for practical issues

regarding the quality of peer tutor performance in APE. The PTEI marks a milestone by providing a quantitative measure to evaluate peer tutor performance in APE. There is not one all-encompassing model to observe and record all aspects of peer tutor performance during APE. The PTEI offers a potential solution for APE practitioners to use to inform their teaching practice and potentially improve their elementary school peer tutor programs.

Table 2.1: Content validity steps (Lynn, 1986)

Stage 1 – Development Stage
Step 1. Identification of content domain (Category)
Step 2. Selection and element creation (Goals, and Objectives)
Step 3. Instrument formation
Stage 2 – Evaluation – Quantification Stage
Step 4. Expert Evaluation / Quantification of Content Validity of each item
Step 5. Expert Evaluation / Quantification of Content Validity of instrument

Table 2.2: Peer tutor evaluation instrument goals and objectives

Main Categories	Goals	Objectives
Communication	Appropriate engagement and interaction with the peer tutee	Peer tutor will stay within 5 feet of tutee during the session. Peer tutor will ask at least 2 general check-in questions during the session.
	Appropriate use of peer tutee name	Peer tutor will use peer tutee name at least four times during the session.
Prompting	Appropriate verbal direction for tutee to perform task and provide processing time	Peer tutor will use correct verbal prompt to begin practice session. Peer tutor will wait at least 10 seconds before giving another direction or asking a question.
	Appropriate use of APE physical prompt when necessary	Peer tutor will use physical prompt only after verbal prompt and demonstration have been given.
Feedback	Appropriate positive specific performance feedback	Peer tutor will offer at least four instances of positive specific performance feedback during the session.
	Appropriate positive specific behavioral feedback	Peer tutor will offer at least 4 different instances of positive specific behavioral feedback during the session.
	Social interaction behavior	Peer tutor will communicate with the tutee on content not related to instruction (i.e., comments on dress, compliments, high fives, etc.) at least four times in the session.

Table 2.3: Expert reviewers' ratings of content validity of items.

Behavior Category	Mean Likert Rating	Percentage Agreement
Communication	3.95	98.75
Prompting	3.60	90.00
Feedback	3.90	97.50
Instrument	3.88	97.00

Note. Range of Likert scores = 0 – 4.

Peer Tutor Evaluation Instrument (PTEI) - Classroom Setting			
Name:		**Date:**	
Evaluator:		**Period:**	
Tutee Name:			
Scoring			
Items 1,4,5, 6: Score the occurrence of the behavior **Items 2, 3, 7, 8, 9: Score the frequency of the behavior**			
Peer Tutor Evaluation Criteria			**Score:**
Appropriate engagement and interaction with the peer tutee			
1. Peer tutor will stay within 5 feet of tutee during the session		**0 1**	
2. Peer tutor will ask at least 2 general check-in questions during the session		**0 1 2**	
Appropriate use of peer tutee name			
3. Peer tutor will use peer tutee name at least four times during the session		**1 2 3 4**	
Appropriate verbal direction for tutee to perform task and provide processing time			
4. Peer tutor will use correct verbal prompt to begin practice session		**0 1**	
5. Peer tutor will wait at least 10 seconds before giving another direction or asking a question		**0 1**	
Appropriate use of APE physical prompt when necessary			
6. Peer tutor will use physical prompt only after verbal prompt and demonstration have been given		**NA / 1**	
****If no prompt necessary – score should be 0 (overall score minus 1)****			
Appropriate positive specific performance feedback			
7. Peer tutor will offer at least 4 instances of positive specific performance feedback during the session		**1 2 3 4**	
Appropriate positive specific behavioral feedback			
8. Peer tutor will offer at least 4 different instances of positive specific behavioral feedback during the session		**1 2 3 4**	
Social interaction behavior comments			
9. Peer tutor will communicate with the tutee on content not related to instruction (i.e., comments on dress, compliments, high fives, etc.) at least four times in the session		**1 2 3 4**	
Calculate a % out of: 22 or (21 - No Physical Prompt)		**Total:**	
Comments:			

Figure 2.1: Peer tutor evaluation instrument. This instrument was created to evaluate peer tutor performance in adapted physical education.

CHAPTER 3

STUDY 2: EFFECTS OF PEER TUTOR TRAINING ON PEER TUTOR PERFORMANCE IN ADAPTED PHYSICAL EDUCATION

Abstract

Peer tutor performance was evaluated in adapted physical education classes for students with severe disabilities. The study employed a single-subject multiple baseline design with 24 participants; 12 students with disabilities and 12 students from general population who served as peer tutors. The intervention phase was comprised of a single session peer-tutor training and booster sessions prior to each adapted physical education class meeting. The Peer Tutor Evaluation Instrument (PTEI) was used to assess peer tutor performance on 7 goals and 9 objectives. Change in the motor skill performance and activity level of the tutees were measured to evaluate the efficacy of the program on these two parameters. Social validity of the intervention was measured using the Intervention Rating Profile (Martens, Witt, Elliott, & Darveaux, 1985) with the special education teachers of the tutees and the classroom teachers of the peer tutors. The results of this study showed that peer tutor training has a positive effect on peer tutor teaching performance as well as motor skill acquisition and activity level for tutees. Favorable social validity ratings from adapted and general education practitioners were also achieved.

Introduction

Children with severe disabilities including autism spectrum disorder [ASD], intellectual disability, multiple disabilities, other health impairment, traumatic brain injury, etc. have unique needs in educational settings (Block, 2007). They have lower academic skills, deficits in social skills, and lags in gross motor skills compared to typically developing peers (Block, 2007; Dunn & Leitschuh, 2010). These students may require one-on-one instruction for academic skills and assistance with self-help

procedures, making them less likely candidates for inclusion in general education settings. Teacher prompting may be required for response to academic probes and social interaction with typically developing peers.

The number of students with disabilities attending our nation's public schools has grown considerably in the past decade (Cervantes, Lieberman, Magnesio, & Wood, 2013) This growth in enrollment comes at the same time as resources decrease. Budget cuts are hampering educational support for students with disabilities. In addition, the number of paraprofessional support staff has consistently been cut in the past 10 years, leading to further limited support for students with disabilities in the classroom (Cervantes et al., 2013). Moreover, larger class sizes make it difficult to provide individualized education to these students (Cervantes et al., 2013; Snyder & Dillow, 2011).

This lack of resources puts a substantial limitation on adapted physical education (APE) for students with severe disabilities. APE is a direct service that is required by law (PL 94-142, the Education for All Handicapped Children Act, & Individuals with Disabilities Acts - IDEA, 1990, 1997, & 2004) to be provided to children with disabilities. If individually designed physical education is prescribed in a child's IEP, the public agency responsible for the education of that child must provide the services directly or make arrangements for those services to be provided through other public or private programs (IDEA, 2004). The law states that resources must be allocated for teaching APE to students with disabilities; however, these resources continue to be limited, placing a heavy burden on practitioners and sacrificing educational experiences for students.

Studies examining Academic Learning Time-Physical Education (ALT-PE) and

students with disabilities in inclusionary physical education classes suggest that students with disabilities consistently have lower amounts of ALT-PE (e.g., learning time) than their typically developing peers (Knowles, Aufderheide, & McKenzie, 1982; Silverman, Dodds, Placek, Shute, & Rife, 1984). Students with severe disabilities often lose out on learning experiences and practice (e.g., ALT-PE) because they require more support from the physical educator, in the form of specialized instruction and modification to activities, equipment, and/or the environment. If the physical educator is unwilling or does not have the knowledge or resources to adapt their teaching strategies and environment to provide students with severe disabilities with the necessary learning supports, these students will not benefit from being in inclusionary classes. Gross (1991) argues that excessive class sizes and inappropriate equipment lead to the unfortunate circumstance that students with severe disabilities are often overlooked in inclusionary settings. These findings suggest instructional programming can be severely lacking for students with disabilities in integrated or inclusive settings (Webster, 1987; Wiskochil, Lieberman, Houston-Wilson & Peterson, 2007). Hence, it is important to have an appropriate setting for these students to succeed in physical education.

Students with severe disabilities need to experience successful motor engagement just like their typically developing peers (Block, 1994; d'Arripe-Longueville, Gernigon, Huet, Winnykamen, & Cadopi, 2002; Ensergueix & Lafont, 2010). Adapted physical education classes can be an effective setting for these students; however, there are some constraints in this setting as well. Time allocated to APE is often restricted. Most elementary APE classes are only 30 minutes per week (Piatti, Beets, & Combs, 2009). Once transition time and time spent organizing and motivating the students to participate

is subtracted from these 30 minutes (Dunn & Leitschuh, 2010), learning time can be as little as 20 minutes a week. Each additional minute of the class should be spent teaching motor and fitness skills and helping students learn to enjoy physical activity, because it is so important in their lives.

A potential solution to improving student-learning outcomes in physical education settings is the use of peer tutors (Block, 2007; Hodge, Lieberman, & Murata, 2012). Peer-assisted learning strategies (PALS) is an evidenced-based program that includes teaching strategies using peers tutors to support direct instruction (PALS; Rosenshine, 1979). This support may use either i.e., unilateral or i.e., bilateral tutoring (Cohen, 1994). When teaching motor or fitness skills to students with severe disabilities, the most common strategy used is unilateral. In unilateral settings, one student is the peer tutor and the other is the tutee. In bilateral settings, the tutoring / tutee roles can be assigned to either participant (Cohen, 1994).

PALS have been proposed as best practice in pedagogy texts for general education and physical education (e.g., Siedentop & Tannehill, 2000) and as a method of inclusion for students with disabilities (Block, Oberweiser, & Bain, 1995). Breslin and Rudisill (2013) found that PALS has the potential to improve student engagement in learning motor skills as well as increase the number of attempts and the rate of response for students with autism. Although the research base supporting PALS in general physical education is growing, few studies have focused on APE students with severe disabilities in self-contained physical education settings (Ward & Ayvazo, 2005).

Using typically developing peers as tutors to assist students with disabilities in the physical education setting is gaining popularity in the literature (Block, 2007; Hodge,

Lieberman, & Murata, 2012; Lieberman & Houston-Wilson, 2009). d'Arripe-Longueville and colleagues (2002) completed a study where one typically developing peer takes on the role of the tutor, and the student with a disability is the tutee in an aquatics skills unit. In the case of this study, higher skilled peer tutors have been shown to yield better results in tutee skill acquisition than less skilled peer tutors, justifying the need for peer tutor training.

Although several studies have explored this topic in inclusive physical education settings, very few have focused on the self-contained setting in which students with similar abilities are grouped in the same class. The self-contained model necessitates peer tutors to be reverse mainstreamed, or brought in from other classes, in order to have appropriate models for behavior and skill demonstration.

The purpose of this study was threefold. The first purpose was to determine the effects of peer tutor training on peer tutor teaching performance as measured by the Peer Tutor Evaluation Instrument. The second purpose was to determine the effect of peer tutors' support on tutees' improvement of motor performance, and number of steps taken (tutor and tutee) in a self-contained APE class. Third, this study sought to determine the social validity of the intervention based on practitioner feedback.

There were two specific research questions this study sought to answer. 1) Does training peer tutors effect tutoring behaviors in APE (i.e., communication, use of peer tutee name, prompting, feedback and social interaction)? It is hypothesized that peer tutor training will affect peer tutor behaviors. 2) To what extent does training for peer tutors affect the motor skill performance and activity level of tutees (i.e., students with severe disabilities in APE)? It is hypothesized that trained peer tutors will affect motor

skill performance and activity level in APE.

Method

Participants

Twelve participants with severe disabilities, 7 boys and 5 girls, ages 6-11 (M = 9.08, SD = 2.11), were randomly selected from four self-contained classrooms in an elementary school to take part in this study. Each of these participants had a current Individualized Education Plan (IEP) at the time of the study. The participants had a variety of disabilities, including intellectual disability, ASD, other health impairment, multiple disabilities, and/or developmental delay. All participants with disabilities were enrolled in a self-contained APE class for 30 minutes per week.

Twelve typically developing students from the same elementary school were selected to participate in the study by the classroom teacher in each of the four classes (3rd and 6th grades). These peer tutors were 5 boys and 7 girls, ages 9-12 (M = 10.33, SD = 1.44). Peer tutors were released from academic learning sessions, hence classroom teachers selected students to participate based on academic performance, citizenship, and willingness to participate. Student participation was conditional on each peer tutor's academic performance. If at any time the classroom teacher noticed a negative effect on the peer tutor's academic performance, the tutor would be dismissed from the peer tutor program. No peer tutors were dismissed from the program.

Peer tutors were assigned to dyads based on suggestions from the special physical educator (e.g., if the student was known to pull hair he/she was assigned a peer tutor with short hair). These dyads were static throughout the course of the study.

Instrumentation

The three dependent variables in this study were peer tutor teaching performance, tutee performance on selected motor skills, and number of steps taken during the physical education class for both tutor and tutee.

The peer tutor evaluation instrument (PTEI) was used to assess peer tutor teaching performance. Tutors were evaluated based on performance concentrating on 7 goals and 9 objectives. The percentage scores for the criteria were calculated and plotted.

The *Test of Gross Motor Development* – 2 (Ulrich, 2000) was used to evaluate tutee performance on selected motor skills for the tutees. The TGMD-2 is norm referenced and used widely for assessing motor skills of children with and without disabilities ages 3 to 10. Students with disabilities participated in a pretest and posttest assessment. Pretest occurred during the second session of the study. Posttest was administered during the 14th session. One participant (Linn) was in a wheelchair for the pretest and posttest data collection. She was not assessed on the same skills as other ambulatory participants. All pretest and posttest assessments were video recorded and scored immediately following the session.

The number of steps for both peer tutors and tutees were measured using CW-701 Yamax Digi-Walker pedometers. This model is a small plastic box that was clipped to the waistband. The pedometers provided an objective assessment of total steps taken in a set amount of time for each individual (Welk, 2002). Pedometers were placed on the peer tutor and tutee at as close to the same time as possible. Steps were recorded for the entire study session.

One Apple iPad Mini video camera was used to video record each data collection

session. The iPad was held by a research assistant in a inconspicuous position in the physical education setting. Wireless voice recorders (Sony ICD-PX440, Sony ICD-UX533) fitted with lavaliere microphones were used to provide clear audio data of peer-to-peer interaction in a loud physical activity setting. The voice recorder was placed in the peer tutor's pocket and the microphone recorded audio within close proximity (5 feet) of the tutor.

Intervention

Peer tutors participated in a 30-minute training session outside of assigned tutor times. In addition, they were provided 3-5-minute booster sessions prior to each subsequent APE tutoring session. The training protocol developed by Klavina et al. (2008) was adapted as a model for training peer tutors in this study. Tutors were trained in a single session due to significant demands on their academic schedule. These time constraints created substantial issues for scheduling, but are a realistic part of public school education and contribute to the generalizability of the study.

The main group training session focused on learning communication, prompting, and feedback skills to be an effective peer tutor. These skills were appropriate communication, using tutee name, providing specific feedback, using proximity, and a system of least prompts.

Appropriate communication was a topic throughout the training. Tutors learned to speak in a friendly way. They were also encouraged to talk with their tutee, even if they did not get an initial response. Peer tutors were encouraged to use their tutee's name, so the tutee knew their peer tutor was addressing them. They were also asked to try and ask at least two check-in questions of their tutee. Proximity is another aspect of

the training. Peer tutors were shown how to stay with their tutee, but not be overbearing. Peer tutors worked on prompts to help tutees participate in activities through the use of statements rather than questions. They also learned they need to wait before giving another prompt. Appropriate methods and situations to use physical prompts were discusses as well. Specific feedback was discussed and practiced so tutors knew how to give feedback that tutees would understand and potentially respond to.

During the training, peer tutors role played methods of teaching, providing feedback, and prompting tutees. Peer tutors were also asked questions to evaluate comprehension of the main concepts listed above. At the end of the training session, peer tutors completed a quiz to evaluate their recollection of concepts; however, a specific score was not required for participation in the study. The quiz was five questions, each question focused on a specific topic covered in the training (e.g., communication, prompting, feedback, etc.) Scores ranged from 85% to 100%, with a mean of 92.5%.

There were treatment fidelity checks for the training in place. The author followed a script, and completed a checklist to ensure each objective of the training was presented. The author conducted all of the training sessions for the sake of consistency. The author also presented all booster sessions during the study.

The booster sessions were less than 3 minutes in length and were designed as a quick review of skills for peer tutors to remember while tutoring. Each booster session topic was determined after compiling performance scores for peer tutors from the previous week. The objective that scored the lowest mean score for the group from the previous week was the topic of the booster session. The training and booster sessions were designed to offer tutors training in instructional assistance for peers with severe

disabilities in a clear and understandable manner.

Social Validity

The Intervention Rating Profile – Modified (IRP-M) was used to measure special and general education teacher perceptions of the social validity of the peer tutor training intervention. These insights offer important feedback for understanding how cooperating teachers perceived the intervention. Separate IRP-M instruments were created for special education and general education teachers to complete. Both IRP-M instruments consisted of the 15 items from the original Intervention Rating Profile-15 (IRP-15; Martens, Witt, Elliott, & Darveaux, 1985). All IRP-15 items were modified to reflect the intervention (peer tutor training) and setting (special education, general education); however, great care was taken to not change the theme of each item. The 15 items on the IRP-15 have factor loadings from .82 to .95 on a single factor that appears to reflect general acceptability (Witt & Martens, 1983). Reliability of the IRP-15 instrument has been reported to be .98 (Kazdin, 1980).

Procedure

Data Collection

Prior to the study, University of Utah and school district IRB approval were obtained. Approved procedures were followed throughout the course of the study. Consent forms were sent to all participant caregivers prior to data collection. In addition, participants completed assent forms prior to the beginning of the study and were reminded that they can abandon the study at any time.

Data were collected on all participants over the course of the 16 APE sessions.

All sessions were video recorded using a mini iPad. Three randomly selected peer tutors (drawing names out of a container) wore the voice recorder and lavaliere microphone clipped to their shirt during each study session. In order to identify peer tutors for evaluation, the peer tutor stated the class, the date, and an assigned letter that was a designation for their name into the audio recorder before each the data collection session began. Voice recorders were used to collect verbal interaction between the peer tutor and tutee. Audio data were coded from recordings after the sessions were completed for the day. Both the peer tutor and the tutee wore a pedometer attached to their waist. Pedometer data were noted and recorded after each teaching session.

All study sessions took place in a multipurpose room, approximately 50 feet long and 30 feet wide. The room was the same one in which the students participated in APE. There was a schedule of activities for the day posted conspicuously on a wall for reference by the peer tutors, staff, and the teacher throughout the study.

Design and Analyses

The data were collected using a multiple baseline design (Shadish & Cook, 2001). This design allows the researcher to forego returning to baseline, which is useful when skill acquisition is not expected to revert to baseline levels. Instead, the intervention is introduced across time and different participants (O'Neill et al., 2011). Baseline data were collected for three sessions prior to introducing the intervention, which was the peer tutor training to group 1. The intervention was introduced to group 1, while groups 2, 3, and 4 were still in baseline. Once the impact of the intervention was established for group 1, the intervention was then introduced to the remaining groups in the same manner (O'Neill et al., 2011). Successive introduction of the intervention with a

corresponding change in the dependent variable across three different groups at three different points in time establishes experimental control in the multiple baseline design (Horner et al., 2005).

Visual analysis was used to inspect and interpret the data (O'Neill et al., 2011). Apparent changes in data patterns corresponding to the experimental manipulation of the intervention were determined by visual analysis. Noting the variability and trends within the conditions, as well as changes in level and overlap between conditions, is essential to determine the effect of the independent variable on the dependent variables (O'Neill et al., 2011). Variability refers to the range of scores within a condition. Trend refers to the slope of the data within conditions and is referred to as flat, increasing, or decreasing (Wolery & Harris, 1982). Level is defined as "the relative value of the data pattern on the dependent variable. Changes in level represent changes in value of the data series as measured on the dependent variable at the point of intervention" (Wolery & Harris, 1982, p. 447). Overlap refers to the degree to which data from the intervention phase fall within the range of data from the baseline phase. The less overlap between conditions, the smaller the variability within conditions and/or the greater the level change. The greater the change in level, trend, and variability between conditions in the expected direction, the stronger the experimental control (Houston-Wilson et al., 1997). Immediacy of the effect of the intervention, that is how quickly the data change after introduction of the intervention, and the consistency of effects across legs of the baseline are (e.g., across groups) considerations for evaluating the effectiveness of the intervention in multiple baseline data (Horner et al., 2005).

Interobserver Agreement

In order to determine interobserver agreement for the Peer Tutor Evaluation Instrument (PTEI), five (30%) of the class sessions were randomly selected to be reviewed by two independent observers. The research assistants' scores were compared with those of the researcher on an item by item basis. Both observers would have to score each criterion the same way (i.e., if the peer tutor states the tutee name at least four times in the course of the practice session, both observers would have to agree on the number of times the peer tutor said the tutee's name and score a four on that objective). Interobserver agreement was calculated as the number of agreements divided by the number of agreements plus the number of disagreements, multiplied by 100. A predetermined mean of $\geq$ 80% interobserver agreement was deemed acceptable. Analysis included reviewing video and audio recordings to complete the Peer Tutor Evaluation Instrument (PTEI) for each peer tutor participating in the class. Observers were allowed to review recordings as many times as necessary to complete the PTEI. Once the PTEI was completed, scores were compared and interobserver agreement was calculated. The range of interobserver agreement was between 78% and 100% with a mean of 85.6%.

TGMD-2

When determining interobserver agreement for the TGMD-2 pretest and posttest assessments, the research assistant needed to concur that the same number and type of TGMD-2 criteria were present in each trial as the researcher for an agreement to be scored. For example, if the participant completed a trial of the underhand roll and performed all of the criteria except for following through, they would score a three out of

a possible four for that trial. Both observers would have to agree that the same criteria were achieved or missing from the trial. Agreement was calculated as the number of agreements divided by the number of agreements plus the number of disagreements, multiplied by 100. A predetermined mean of ≥ 80% interobserver agreement was deemed acceptable. IOA was calculated for both the pre and posttest assessments. Video recordings for both pre and posttest phases of the TGMD-2 were reviewed by a trained observer. The research assistant reviewed the video and scored the number of criteria present for the skill. Scores were then compared with those of the researcher and IOA was calculated. The range of IOA was 94% to 100% with a mean of 96.7%.

Results

Peer Tutor Performance

Data were collected in a multiple baseline across participants design (Kazdin, 1982) to evaluate peer tutor teaching performance as measured by the PTEI before (baseline) and after Peer Tutor Training. There were four separate classrooms in which peer tutors participated. Individual peer tutor performance for each group is presented in Figure 3.1. The mean and standard deviation of performance scores during intervention and baseline phases are shown in Table 3.1. Intervention was introduced to each group after baseline data points were determined to be stable. In this study, intervention for group 1 was introduced after three baseline data points, for group 2 after six baseline data points, for group 3 after nine baseline data points, and for group 4 after 10 baseline data points.

Group 1

The three tutors in Group 1 were from a 3rd-grade classroom. These tutors worked with tutees from a K-3 Functional Academics classroom. Tutees from this classroom were verbal, compliant, and had moderate to good attending skills. The three tutees involved in the study were diagnosed with intellectual disability (n = 1) and other health impairment (n = 2). All three tutees have rarely had any behavior difficulties in APE, and all seemed to enjoy learning and being active in the setting. One tutor/tutee dyad from this group was removed from the study after week 12 due to the tutee moving out of state (Lisa, Rox). Data for this tutor were included, but were noted as incomplete.

Percentage of criteria met on the PTEI for Lisa, a peer tutor, in baseline were collected for 3 weeks and were stable (M = 8%), without trend or variability (see Figure 3.1). Intervention phase data were collected with Lisa for 8 sessions (M = 56.04%; SD = 7.7%). When comparing means, a substantial level change after intervention is apparent. Trend in intervention was relatively flat based on the range of scores; however, there was a decrease in performance in week 5. After spring break, Lisa's performance improved (week 8) and maintained a similar level until the dyad was removed from the study. There was some variability in the data, although there was no overlap between phases.

Percentage of criteria met on the PTEI for April in baseline were collected for 2 weeks and were stable (M = 18%; SD = 11.5%); trend for baseline was increasing, and there was some variability in the data. A limitation of the baseline data for April is that she was absent during the final baseline week. This final data point would have provided a clearer idea of trend for this tutor. Intervention phase data were collected for 13 weeks (M = 71%, SD = 11.5%). Trend in intervention showed a moderate increase. The change

in level was immediate and large following intervention, when comparing means between phases. There was some variability in the data, but no overlap between phases. April's tutee missed these 3 days of the study (weeks 5, 7, and 11; designated by an X); consequently, she was asked to work with a different tutee these days. These points represent slight variability in the performance scores.

Percentage of criteria met on the PTEI for Sam in baseline were collected for 3 weeks and were stable (M = 14.0%, SD = 2.2%), trend for baseline was level, and there was slight variability in the data. Intervention phase data were collected for 13 weeks (M = 56.2%, SD = 12.5%). When comparing means, it was determined there was a significant level change in performance after intervention. Trend in intervention increased; however, there was a decrease in performance in week 5. After spring break, Sam's performance improved slightly (week 6) and continued improving to the end of the study. There was some variability in the intervention data, although there is no overlap with baseline.

Group 1 peer tutor performance data show a substantial increase in performance scores after intervention when comparing means across phases for each participant. There is an immediacy of treatment effect (e.g., significant increase in performance scores immediately after intervention) present for all tutors in this group as well. Tutor scores do not show any overlap with baseline data, suggesting positive effects of peer tutor training on peer tutor performance.

Group 2

The 3 peer tutors in Group 2 were from a 6th-grade classroom. These tutors worked with tutees from a 4th-6th-grade Functional Academics classroom. The three

tutees in group 3 were diagnosed with intellectual disability *(n* = 1), other health impairment *(n* = 1), and autism *(n* = 1). Two of the three tutees have rarely had any behavior difficulties in SPE, and seem to enjoy learning and being active in the setting. They were interactive, verbal, somewhat higher skilled in SPE, compliant, and had excellent attending skills. The third tutee (Rex) was less interactive, had limited verbal skills, and sometimes refused to attend the skill sessions. He would walk or run around the space while his tutor prompted him to join the group; however, a reinforcement system that was introduced by the special education teacher in week 6 (i.e., stamps, edibles) improved participation and decreased evasive behaviors.

Percentage of criteria met on the PTEI for Wayne, a peer tutor, in baseline were collected for 6 weeks and were less stable (*M* = 30.1%, *SD* = 7.5%). Trend for baseline increased slightly, but with more variability than other peer tutors in this group. Intervention phase data were collected for 10 weeks (*M* = 68.7%, *SD* = 6.7%). When comparing means across phases, Wayne demonstrated a large, immediate increase after intervention. Trend in intervention decreased over time; however, there was no overlap between baseline and intervention. Wayne's performance reached a peak at week 9 (80%) and trended downward through the end of the study. There was less variability in the intervention data when compared to baseline.

Percentage of criteria met on the PTEI for Pre in baseline were collected for 4 weeks and were relatively stable (*M* = 26.0%, *SD* = 6.2%); trend for baseline was slightly decreasing. Pre was absent during the third and sixth baseline sessions. Secondarily, Pre was the tutor assigned to work with Rex, the tutee who participated inconsistently in APE sessions. Intervention phase data were collected for 10 sessions (*M* = 71.6%, *SD* =

4.4%). Trend in intervention showed a slight decrease with a more substantial drop for the last session (although still above baseline levels). The change in level was immediate and large following intervention, when comparing means between phases. There was some variability in the data that occurred throughout intervention phase.

Percentage of criteria met on the PTEI for Dan in baseline were collected for 6 weeks and were moderately stable (M = 26.7%, SD = 6.5%); trend for baseline increased slightly. Intervention phase data were collected for 10 weeks (M = 71.6%, SD = 14.5%). Trend in intervention showed a clear increase. There was some variability in the intervention data with point 13 showing a decrease in score, although there is no overlap with baseline.

Group 2 data showed a substantial increase in performance scores after intervention. There was an immediacy of treatment effect present for all tutors in this group as well. Tutor scores did not show any overlap with baseline data.

Group 3

Tutors in Group 3 were from a 3rd-grade classroom. These tutors worked with tutees from a K-3 functional life skills classroom. Tutees from this classroom were sometimes unpredictable, exhibited aggressive behavior (e.g., hair pulling, hitting, biting, etc.), were noncompliant, nonverbal, and/or emotionally unstable (e.g., crying, yelling, etc.). Each tutee in this class presented a unique challenge for the peer tutors. Erik exhibited aggressive behaviors (e.g., hitting, pushing, kicking, hair pulling, etc.) toward his peers; however, he never showed aggression toward his peer tutor. Susy had limited verbal skills, and was sometimes noncompliant. Mit displayed high rates of avoidance behaviors (e.g., crawling under tables, running out any unattended exit door). The 3

tutees involved in the study were diagnosed with developmental delay. These 3 tutees had difficulty transitioning, following directions, participating, and having a calm body in APE. Each tutee had a behavior management plan. There were no fewer than five adults present at each class to support tutors and tutees during this study.

Percentage of criteria met on the PTEI for Misty in baseline were collected for 9 weeks and were quite stable (M = 35.3%, SD = 5.0%); trend for baseline showed a slight decrease. Intervention phase data were collected for 5 weeks (M = 41.7%, SD = 5.6%). Trend in intervention was slightly upward, but with more variability. Session 13 data are missing due to academic testing. Data for Misty showed a slight overlap between baseline phase and intervention phase for one data point (12). There were larger amounts of variability in the intervention data than were present for the other peer tutors in this group.

Percentage of criteria met on the PTEI for Dax in baseline were collected for 8 weeks and were relatively stable (M = 30%, SD = 11.9%); trend for baseline is slightly increasing. A limitation of the baseline data for Dax was that she was absent during the ninth baseline session. Intervention phase data were collected for 5 weeks (M = 61%, SD = 6.3%). When comparing means, there is a significant increase in performance. Trend in intervention showed an increase through the completion of the study. There was a significant level change after intervention. There was some variability occurring throughout intervention phase; however, no overlap with baseline phase data.

Percentage of criteria met on the PTEI for Evan in baseline were collected for 8 weeks and were relatively stable (M = 36.1%, SD = 9.6%); trend for baseline was marginally increasing. Intervention phase data were collected for 5 weeks (M = 63.5%,

SD = 3.5%). Trend in intervention increased, but with substantial variability in the data. Overlap occurred between baseline phase and intervention phase (11).

Group 3 data showed a marginal increase in performance scores after intervention. There is an immediacy of treatment present for all tutors, except for Evan. Evan's initial score after intervention showed overlap with baseline phase.

Group 4

The peer tutors in Group 4 were from a 6th-grade classroom. These tutors worked with tutees from a 4-6 functional life skills classroom. Tutees from this classroom were sometimes unpredictable, exhibited aggressive behavior (e.g., hair pulling, hitting, biting, etc.), were noncompliant, nonverbal, and/or emotionally unstable (e.g., crying, yelling, etc.). Each tutee in this class presented a unique challenge for the peer tutors. Lucy has limited mobility and uses a wheelchair or walker. Avery was sometimes emotional (e.g., crying, yelling or moody), as well as noncompliant, and exhibited evasive behaviors. Gus was sometimes noncompliant, violated personal space (e.g., inappropriate hugging, kissing, grabbing arms), and verbalized derogatory comments (e.g., swearing, name calling) toward peers, tutors, and staff. The three tutees involved in the study were diagnosed with other health impairment (n = 1), intellectual delay (n = 2). These three tutees sometimes struggled with transitioning, following directions, and attending in APE. Each tutee has a behavior management plan created by the classroom teacher. During this study, there were no fewer than four adults present at each class to support tutors and tutees.

Percentage of criteria met on the PTEI for Tim in baseline were collected for 7 weeks and were highly unstable (M = 35.3%, SD = 11.3%), trend for baseline showed a

slight increase. Tim was absent 5 out of 12 sessions during baseline. Intervention data were collected for 4 weeks (M = 53.6%, range, SD = 2.8%). There is a substantial level change when means are compared. Trend in intervention was flat, but with much less variability. There was significantly less variability in the intervention data when compared to baseline.

Percentage of criteria met on the PTEI for Abby in baseline were collected for 12 weeks and were relatively stable (M = 30.1%, SD = 4.6%), trend for baseline is flat. Intervention data were collected for 4 weeks (M = 67.6%, SD = 4.5%). There is a substantial level change between baseline and intervention when means are compared. Trend in intervention showed a slight increase through the completion of the study. There was slight variability throughout intervention, but no overlap between phases.

Percentage of criteria met on the PTEI for Mel in baseline were collected for 12 weeks and were reasonably stable (M = 36%, SD = 8.0%), trend for baseline increased slightly. Intervention phase data were collected for 4 weeks (M = 84.3%, SD = 6.0%). There is a substantial level change between baseline and intervention when means are compared. Trend in intervention decreased; however, there was no overlap with baseline data. There was moderate variability in the intervention data.

Percentage of criteria met on the PTEI for Group 4 showed a marginal increase in performance scores after intervention, with Mel being an exception. Her performance scores improved immediately and substantially after intervention; however, these scores for intervention trended downward, suggesting potential issues with resiliency of the intervention.

Overall Evaluation of Peer Tutors' Teaching Performance

Evaluating the overall teaching performance of across all 4 groups provides an opportunity to consider replication. All groups showed relatively stable baseline data, moderate amounts of variability, and immediate, large increases in performance after the introduction of the intervention. These patterns support fairly strong experimental control (Houston-Wilson et al., 1997). Moreover, there is evidence supporting strong immediacy of effect of the intervention and very little overlap (two data points) between baseline and intervention phases. Most peer tutor performance scores had flat or increasing slopes in intervention as well.

A critical element for considering replication is the consistency of effects across those groups. Groups 1 and 2 show much higher effects of intervention than groups 3 and 4. Groups 1 and 2 were functional academics classes with tutees who were more compliant, verbal, interactive, and higher skilled. Tutees in groups 3 and 4 were from functional life skills classes and were less predictable, potentially aggressive, noncompliant, less interactive, and lower skilled. Attention toward within group effects of the intervention supports the idea that the effect of the treatment was mediated by the tutees in some way. When groups 1 and 2 are compared, the change in level was similar. The same was true for groups 3 and 4. There was a level change for all groups; however, it was not consistent across all groups. The change was more consistent between groups of tutees in functional academics and functional life skills classes.

Tutee Step Count

Tutee step count data were recorded to identify potential correlation between peer tutor training and tutee step counts, that is, would a trained peer tutor be more successful

at instigating more activity from a tutee than an untrained peer tutor? Step counts were recorded using a pedometer and were collected at the conclusion of each session. Table 3.2 presents the mean and standard deviation of step counts for each tutee during baseline and intervention. Figure 3.2 displays the tutee step counts in a linear format with each point representing the number of step counts per session. Visual inspection of these data utilizing the same methods detailed above (variability, trend, overlap, immediacy of effect) reveal several note-worthy elements.

Group 1

Group 1 tutees were from a k-3 functional academics classroom. These tutees worked with tutors from a 3rd-grade classroom. Tutees in this class were compliant with wearing the pedometers and did not require any further support from teachers or tutors. Most tutees were interested in the number of steps they acquired during a session. Tutors and tutees repeatedly engaged in contests to see who would have the highest number of steps. All tutees observed the rules of wearing the pedometers and none were observed breaking them.

Step count data for Rox in baseline were collected for 3 weeks and were unstable ($M = 1107$, $SD = 477$); trend for baseline increased. Intervention phase data were collected for only 4 sessions ($M = 1244$, $SD = 320$) because Rox was exited from the study when she moved away. Trend in intervention was slightly increasing, but with less variability in sessions 8 through 11. The data showed a level change (137 steps) when comparing means; hence, there does appear to be an immediacy of effect on step counts due to the peer tutor training.

Step count data for Louise in baseline were collected for 2 weeks and were

unstable (M = 1147, SD = 385); trend for baseline increased, with only two data points for interpretation. Intervention data were collected for 7 weeks (M = 1180, SD = 290). Trend in intervention slightly increased through the completion of the study. When means are compared, there was a minimal level change (33 steps) after intervention. There was significant variability occurring throughout intervention phase as presented by the range of step counts.

Step count data for Pat in baseline were collected for 3 sessions and were reasonably stable (M = 1195, SD = 136); trend for baseline was decreasing. Intervention data were collected for 12 sessions (M = 1725, SD = 253). Trend in intervention increased (530 steps) when comparing means. There was modest variability in the intervention data.

Step count data for tutees in Group 1 showed a marginal increase in step counts after intervention, with Pat being an exception. His step counts increased 3.5 times that of his peers after intervention. There is a modest immediacy of treatment present for all tutees.

Group 2

Group 2 tutees were from a 4th-6th-grade functional academics classroom. These tutees worked with tutors from a 6th-grade classroom. Tutees from this class wore the pedometers and were interested in the number of steps they acquired during a session. All tutees observed the rules of wearing the pedometers.

Step count data for Rex in baseline were collected for 4 weeks and were highly unstable (M = 1030, SD = 253); baseline trend increased. Intervention data were collected for 8 weeks (M = 1243, SD = 124); intervention trend decreased, but with much less

variability than baseline. The data showed a change in level (213 steps) when comparing the baseline and intervention means.

Step count data for Wyatt in baseline were collected for 4 weeks and were highly variable (M=1075, SD = 804); trend for baseline increased; however, there were only four data points for interpretation, instead of six possible. Intervention data were collected for 8 sessions (M = 1487, SD = 205). Trend in intervention increased through the completion of the study. When comparing means, there was a substantial level change (412 steps) after intervention. There was significantly less variability in intervention as represented by the range of step counts.

Step count data for Sven in baseline were collected for 6 weeks and were highly unstable (M = 888, SD = 365); baseline trend was flat. Intervention phase data were collected for 8 weeks (M = 1011, SD = 357). The level in intervention increased when comparing means (123 steps). There was some variability in the intervention data.

Group 2 step count data showed a marginal increase after intervention when means are compared; however, there were significant amounts of variability in the data.

Group 3

Group 3 tutees were from a K-3 functional life skills classroom. These tutees worked with tutors from a 3rd-grade classroom. Two tutees from this class exhibited minimal compliance with wearing the pedometers. Paraeducators attempted to slip the instrument on the student's waist, and were successful six sessions with Mel, and seven sessions with Eddy. If these students discovered the pedometer, they would remove it and throw it across the room. Susie did not exhibit any adverse behaviors to wearing the pedometer. These three tutees did not show any interest in the step count data recorded

on the pedometer.

Step count data for Mel in baseline were collected for 6 of 9 weeks and were highly unstable ($M = 470$, $SD = 277$); baseline trend decreased. Intervention data were collected for 3 sessions ($M=369$, $SD = 166$) because Mel refused to wear the pedometer in several sessions. Trend in intervention increased slightly; however, there is significant overlap between baseline and intervention phases. When comparing means, the data showed a slight decrease in level from baseline to intervention (-101 steps).

Step count data for Susie in baseline were collected for 8 weeks and are not stable ($M = 1054$, $SD = 271$); trend for baseline decreases slightly. Intervention phase data were collected for 4 sessions ($M = 1228$, $SD = 229$). The level in intervention increased slightly over baseline (174 steps). There was variability in the intervention data.

Step count data for Eddy in baseline were collected for 7 weeks and are unstable ($M = 601$, $SD = 422$). Trend is difficult to interpret due to the extreme variability of the baseline data. There were several sessions (3, 4, 9, 10, 13) where Eddy either threw the pedometer multiple times, or refused to wear it (6, 15). Intervention phase data were collected for 4 weeks ($M = 450$, $SD = 188$). Trend in intervention decreased through the completion of the study. When comparing means, a slight level change (-151 steps) between baseline and intervention was determined. There was some variability occurring throughout intervention.

Step count data for Group 3 showed an overall decrease in step counts after intervention, with Susie being an exception. The mean for her step counts presented the only increase in her group after intervention. Consideration for level change between baseline and intervention phases showed a small increase for Susie; however, these data

decreased for both Mel and Eddy.

Group 4

Group 4 tutees were from a 4-6 functional life skills classroom. These tutees worked with tutors from a 6th-grade classroom. Two tutees in from this class exhibited moderate compliance with wearing the pedometers, once a reinforcement protocol was in place. Gus was less compliant; however, compliance increased once he was told it was a super hero spy camera. Pedometer compliance was mood dependent for Gus. If he discovered it, he would throw it across the room. These three tutees did not show any interest in the pedometer or step counts, with the exception of Avery. She repeatedly attempted to shake the pedometer to increase her steps and "beat" the count of her tutor.

Step count data for Gus in baseline were collected for 9 of 13 weeks and were highly unstable (M = 885, SD = 334), trend for baseline decreased moderately. Intervention phase data were collected for 2 weeks (M = 774, SD = 356). Trend in intervention increased; however, there was significant overlap between baseline and intervention phases. The data showed a slight decrease in level when comparing means (-111 steps).

Step count data for Linn in baseline were collected for 12 weeks and were relatively stable (M = 203, SD = 305) trend for baseline increased. These data for Linn showed low levels of step counts, aside from sessions 4 and 11, because Linn chose to use her wheelchair rather than her walker in APE. When Linn used her wheelchair, the pedometer was placed on her hip to attempt to record any hip movement in an effort to move her legs in a stepping motion. Intervention phase data were collected for 3 sessions (M = 202, SD = 207). Level in intervention was flat (-1 step) when means were

compared. There was variability in the intervention data due to another session (15) where she chose to use her walker.

Step count data for Avery in baseline were collected for 11 weeks and were moderately stable (M = 1199, SD = 291). Trend increased moderately. Intervention phase data were collected for 3 weeks (M = 1182, SD = 229). Trend in intervention showed a moderate increase. There was a level change when comparing means (-17 steps) after intervention, indicating substantial overlap in the data. Moderate variability occurred in the data throughout intervention.

Group 4 data showed an overall decrease in step counts after intervention when comparing means (see Table 3.2). Level change decreased between baseline and intervention phases for the three tutees in this group.

Tutee Skill Acquisition

TGMD-2 (Ulrich, 2002) pretest and posttest data were collected to evaluate skill acquisition for students with severe disabilities. Predetermined skills of gallop, skip, overhand throw, and underhand roll were selected as skills that students had the greatest potential for improvement based on pretest scores. These skills are rated as more difficult according to the TGMD-2 manual (Ulrich, 2002). We decided to assess a more difficult set of skills to preclude a ceiling effect. Pretest and posttest mean performance scores are summarized in Table 2.3, including percent change and percent mastery.

Pretest and posttest means were analyzed using SPSS with a paired samples t-test (Field, 2013) in order to determine if there was a statistically significant difference in pretest and posttest scores for all groups. Data were combined across groups for each skill. Pretest and posttest scores for each skill were compared.

There was a significant change for participant TGMD-2 skip scores for pretest (M = 1.08, SD = .956) and posttest (M = 2.08, SD = 1.40). This difference was significant $t(3) = -4.19$, $p = .025$, and represented a large effect size, $d = .83$. There was also a significant change for TGMD-2 gallop scores for pretest (M = 1.00, SD = .670) and posttest (M = 1.62, SD = .644). This difference was significant $t(3) = -4.42$, $p = .021$, and represented a large effect size, $d = .91$. TGMD-2 scores for underhand roll and overhand throw did not change significantly from pretest to posttest due to high rates of variance in the scores for participants.

Social Validity

Three special education teachers completed the IRP-M SPED at the conclusion of the study. Scores were compiled to identify overall favorability to the intervention from a practitioner's perspective. The scores from the IRP-M SPED ranged from 4 (slightly agree) to 6 (strongly agree) with a mean score of 5. The item scoring the lowest was item #10, "The teacher expectations for this intervention were clearly described."

Three general education teachers completed the IRP-M GENED at the conclusion of the study. The scores from the IRP-M GENED ranged from 5 (agree) to 6 (strongly agree) with a mean score of 5 from the general education teachers.

Discussion

This study evaluated the effects of a peer tutor training intervention on peer tutor teaching performance. The results from the multiple baseline design showed an overall positive effect of the training program on peer tutor teaching performance. The relative improvement in teaching performance of all peer tutors after intervention supports the

prediction that peer tutor training is an important aspect of a successful peer tutor program in APE. These improvements compared favorably with previous studies (Houston-Wilson, 1997; Klavina, 2001), showing that elementary age peer tutors can successfully teach skills to students with disabilities in self-contained SPE classes, or an inclusive PE class.

Also evaluated in this study were the effects of a peer tutor training intervention on tutee skill acquisition. The results of this study demonstrated a significant improvement scores on the TGMD-2 (Ulrich, 2002) for two of the four motor skills. A substantial body of literature supports using peer tutors to teach students in APE (Barfield et al., 1998; Block, 2007, 1994; Block et al., 1995) with a focus on improving motor skills.

This study also supports the need for tutors to engage in training (d'Arripe-Longueville et al., 2002; Houston-Wilson, 1997; Klavina, 2001) to be as successful as possible engaging with and teaching students with severe disabilities in SPE. Other studies have proposed longer training sessions for peer tutors working with students with severe disabilities (Klavina, 2002; Long et al., 1980). This study supports findings from Houston-Wilson (1997) by showing that elementary-aged students can be taught to assist students in relatively simple terms and a short amount of time. This realization is important because of the increasing time constraints students face in elementary school (e.g., standardized testing, specific academic support, field trips, educational fairs, etc.). These findings show that if a small amount of time can be set aside for training and booster sessions provided for a few minutes before each class, there is a great potential for peer tutor teaching performance improvement.

Conclusions

The students with severe disabilities in this study demonstrated delayed motor skills relative to their typically developing peers. The students demonstrated a mean TGMD-2 score of 22% for locomotor skills and 28% for object control skills. Generally, these scores mean students can perform one out of four criteria for demonstrating a motor or object control skill. This level of performance is significantly lower than typically developing peers, which presents a strong case for as much intervention and support for these students as possible in the APE setting.

The peer tutor training program and PTEI assessment were developed for this study and were implemented to create an opportunity to train and evaluate peer tutors. The peer tutor training program protocol focused on teaching tutors how to establish a rapport with students by being in close proximity to the tutee and by talking and interacting with the tutee. It also focused on prompting with statements as well as appropriate physical prompts. Lastly, providing specific feedback relating to performance and behavior was discussed. For example prior to training, peer tutors would simply say "good job," or not say anything to their tutee for the entirety of the class. After training, peer tutors would give specific performance feedback such as, "Sven I like how you kicked the ball at the target. This time try to kick it with the inside of your foot." At this point the peer tutor was touching the inside of the tutee's foot and showing the tutee where to contact the ball. Training peer tutors that it is appropriate to use a physical prompt if necessary helps tutees hear and feel the correct elements of a skill. Specific feedback from peer tutors seems to be more meaningful coming from a peer tutor who is trained to establish rapport with their tutee. This type of attention is difficult for

classroom teachers to give because of broad ranges of disability, large class sizes, and significant time constraints. Thus, trained peer tutors can offer much needed support to a learning environment for students with severe disabilities in APE.

Future studies should examine the effect of a more lengthy peer tutor training (3-4 sessions) to improve peer tutor teaching performance. Studies might also consider how the attitudes of peer tutors are affected by training and by their interaction with students with severe disabilities. For example, do peer tutors feel more confident, less nervous, happier, and more empathetic, while working with their tutee after completing the peer tutor training intervention? Further research could also consider another school setting with different age groups, such as junior high or high school, to begin examining how to best train peer tutors of different ages. Finally, similar studies could be repeated over a longer period of time to determine if peer tutor's teaching skills continue to improve over time and if tutees continue to improve their motor skills, and number of steps taken during the session.

Table 3.1: Peer tutor teacher performance scores – PTEI: mean and standard deviation.

Group	Baseline		Intervention	
	M	*SD*	*M*	*SD*
Group 1				
Sam	14.0	2.2	56.2	12.5
April	18.0	2.8	71.0	11.5
Lisa	8.0	0.0	56.0	7.7
Group 2				
Dan	26.7	6.5	71.6	14.5
Pre	26.0	6.2	71.0	4.4
Wayne	30.1	7.5	68.7	6.7
Group 3				
Evan	36.1	9.6	63.5	3.5
Dax	30.1	11.9	61.5	6.3
Misty	35.3	5.0	41.7	5.6
Group 4				
Mel	36.1	8.0	84.3	6.0
Abby	30.1	4.6	67.6	4.5
Tim	35.3	11.3	53.6	2.8

Note: Mean scores are percentage scores out of 100% possible

Table 3.2: Tutee step count – mean and standard deviation.

Group	Baseline		Intervention	
	M	*SD*	*M*	*SD*
Group 1				
Pat	1195	136	1693	253
Louise	1147	385	1302	290
Rox	1107	477	1244	320
Group 2				
Sven	888	365	1011	357
Wyatt	1075	804	1487	205
Rex	1030	253	1243	124
Group 3				
Eddy	601	422	450	188
Susie	1054	271	1228	229
Mel	470	277	369	166
Group 4				
Avery	1199	291	1182	229
Linn	203	305	202	207
Gus	885	334	774	356

Table 3.3: TGMD-2 Mean scores for pretest and posttest including percent change and percent mastery.

Underhand Roll	Pretest *M*	Posttest *M*	% Change	%Mastery
Group 1	2.33	3	0	75
Group 2	0.66	3	200	75
Group 3	0.33	0.66	33	16.6
Group 4	1.33	1.66	16.66	41.6
Overhand Throw				
Group 1	4	3.66	-8.33	91.66
Group 2	0	1.66	166	41.6
Group 3	0.33	0.66	33	16.6
Group 4	0.33	1.33	100	33.3
Skip				
Group 1	2.33	4	25	100
Group 2	1	2	100	50
Group 3	0	0.66	66.6	16.6
Group 4	1	1.66	66.6	41.6
Gallop				
Group 1	2	2.5	25	62.5
Group 2	0.66	1.66	77.6	55.3
Group 3	0.66	1.33	44.3	44.3
Group 4	0.66	1	33.3	30.3

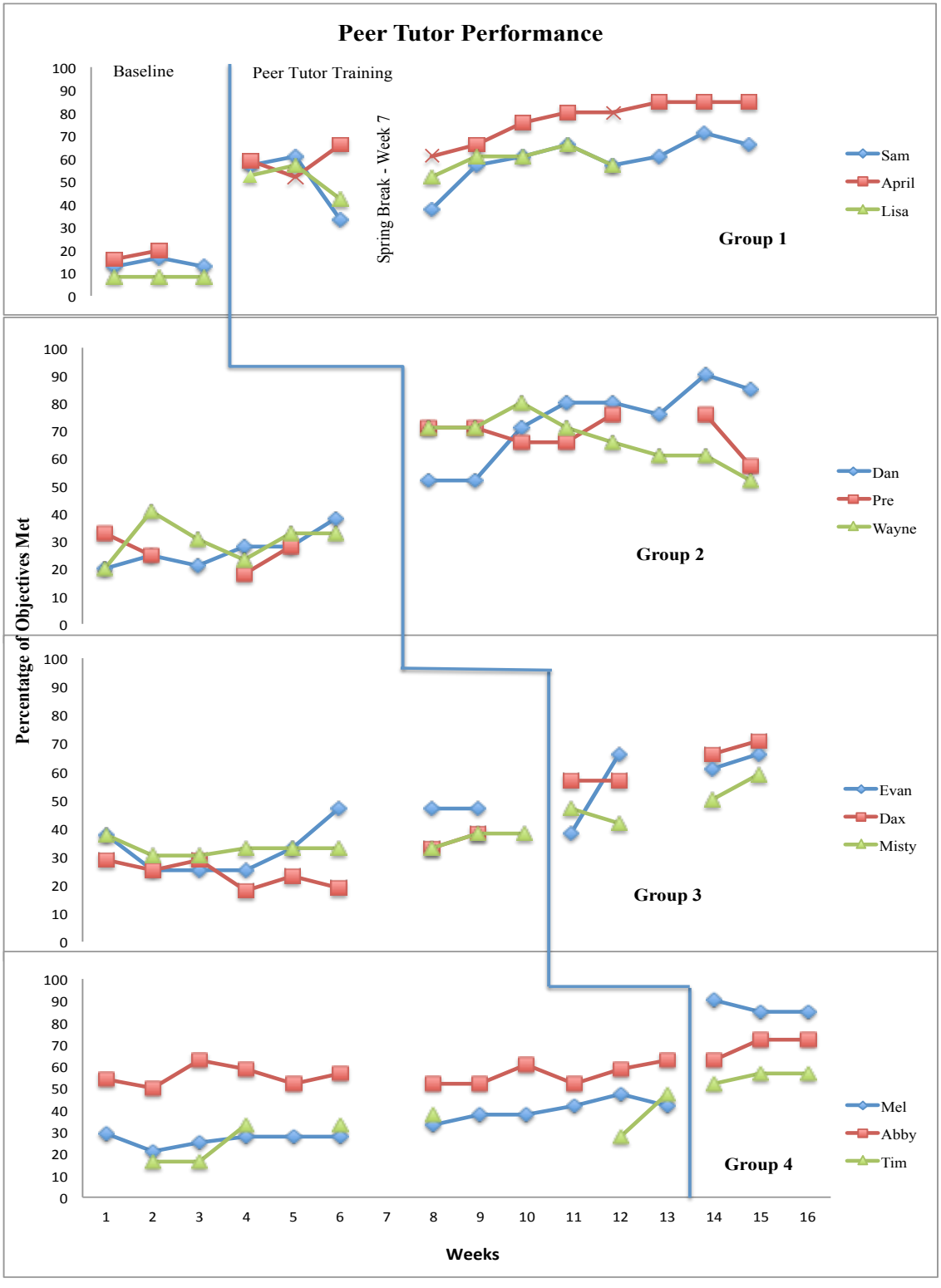

Figure 3.1: Peer tutor performance scores on the Peer Tutor Evaluation Instrument (PTEI).

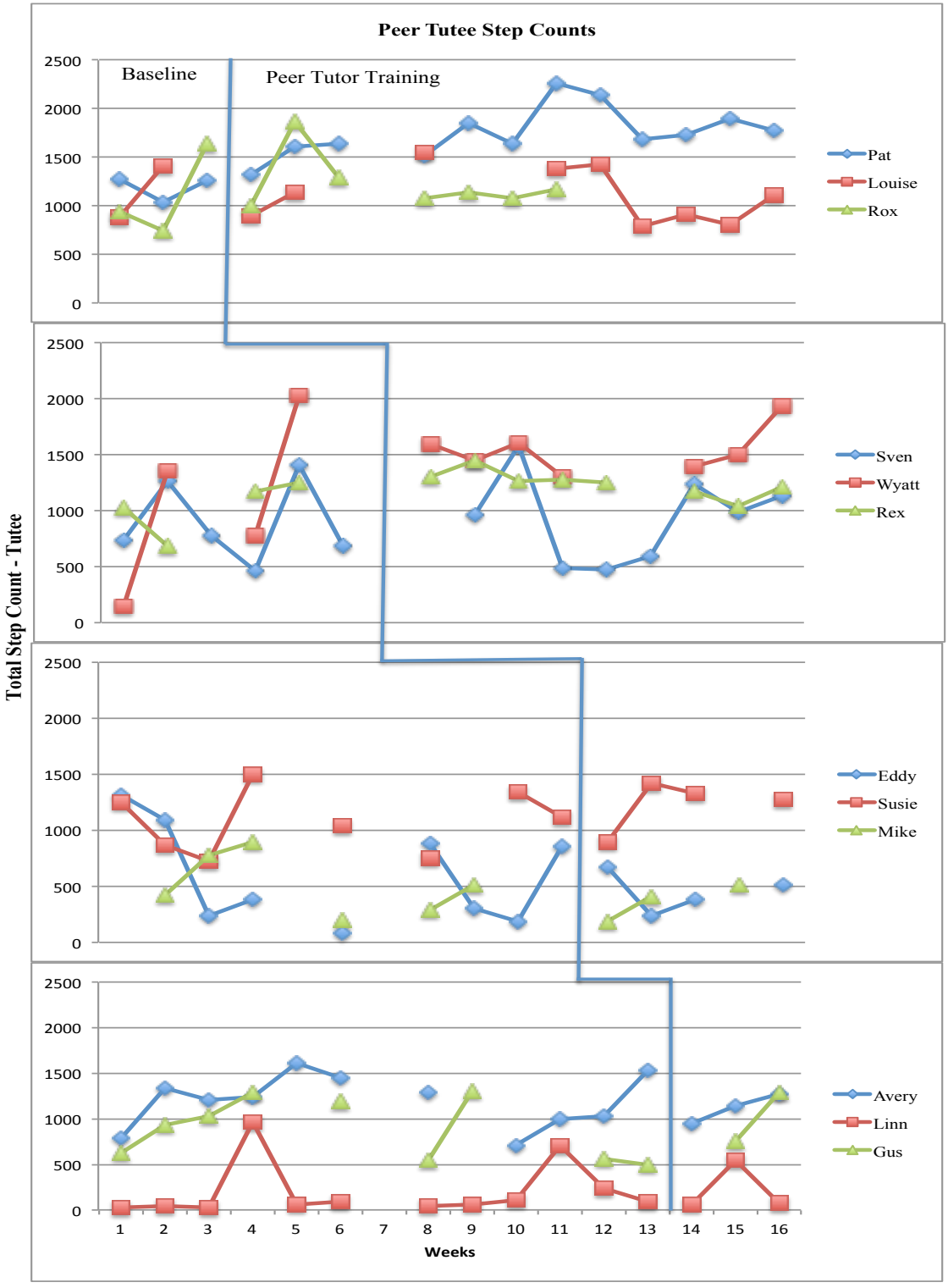

Figure 3.2: Peer tutee step counts. This figure illustrates the step counts for tutees in each session.

CHAPTER 4

STUDY 3: EVALUATION OF PEER TUTOR ATTITUDES TOWARD STUDENTS WITH SEVERE DISABILITIES IN ADAPTED PHYSICAL EDUCATION

Abstract

This study investigated the change in peer tutor attitude toward students with severe disabilities during the participation in a peer tutor program in an adapted physical education setting. Peer tutors (5 males, 7 females) were trained in a 30-minute session. This training was coupled with 3-minute booster sessions prior to each class. Peer tutor attitude data were collected using a smiley face instrument (3-point Likert scale) during a 15-session multiple baseline design to discern change in peer tutor attitude before and after participation in the peer tutor training program. The measure was implemented after each session during the course of this study. A Wilcoxson matched pairs test revealed that there was a small improvement in peer tutor attitude toward students with severe disabilities. It was noted that the level of attitude began at a relatively high rate, suggesting a ceiling effect may have been present. Results of the study, as well as anecdotal evidence, support the hypothesis that peer tutor attitudes would change following participation in a peer tutor program. Moreover, anecdotal and qualitative evidence support Allport's contact theory (1954) that attitudes will improve with increased contact.

Introduction

Limited research has been conducted on the attitudes of peer tutors toward peers with disabilities in the adapted physical education setting (De Boer, Pijl, & Minnaert, 2012); however, attitudes of peers can be one of the most important variables in the success of a peer tutor dyad (Lieberman, 2012). Not all peer tutors have positive attitudes toward their peers with disabilities; however, training for peer tutors can improve tutoring skills (Klavina, 2008) and attitudes toward peers with disabilities

(Slininger, Sherrill, & Jankowski, 2000).

Attitude is a complex construct with several definitions (Ajzen & Fishbein, 1980; Antonak & Livneh, 1988; Tripp & Sherrill, 1991). The construct of attitude has generally been restricted to an individual's evaluation of a psychological experience (e.g., experience with other individuals; Ajzen & Fishbein, 1980). Some theorists consider attitude as being multidimensional and focused on the importance of measuring cognitive, affective, and behavioral aspects of attitude. According to Triandis (1971) "An attitude is an idea (cognitive component) charged with emotion (affective component) which predisposes (cognitive component) a class of actions (behavioral component) to a particular class of social situations" (p. 199). Sherrill (1998) streamlines the concepts of Triandis (1971) by describing attitude as: "… an enduring set of beliefs charged with emotion that predisposes a person to certain kinds of behaviors" (p. 7). Discussion of definitions for attitude illustrates the complexity of the construct researchers grapple with when studying this topic.

Although definitions may vary, there seems to be agreement in the literature that attitude research should be based a specific set of six characteristics. Attitudes are 1) learned through experience and interaction with other people; 2) complex, 3) resistant to change and relatively stable; 4) have a specific social object as a referent, 5) vary in their quantity and quality, possessing different degrees of motivating force (intensity, strength), and direction (toward, against, away from the attitude referent); and 6) are manifested behaviorally via predisposition to act in a specific way when the individual encounters the attitude referent (Slininger, Sherrill, & Jankowski, 2000). These six characteristics provided the framework for this study of peer tutor attitudes toward their

tutees with severe cognitive disabilities in an adapted physical education setting (Slininger et al., 2000).

Slininger et al. (2000) suggest contact theory as a guiding principal for studying attitudes in APE. Contact theory provides a lens through which to view the previously mentioned characteristics of attitude. This study assumes that the tutees who are substantially different from the peer tutor group may be marginalized in the educational setting (Dunn, & Leitschuh, 2000; Henderson, Lavay, & French, 2002; Slininger et al., 2000). Marginalization can result from various degrees of negative behavior, avoidance, discrimination, physical attack, and/or bullying (Bourke, & Burgman, 2010).

Contact theory provides a framework from which to study the relationships between peer tutor and tutee. Contact theory suggests that prejudice and discrimination toward a minority group will be reduced when contact between individuals meets four criteria: 1) there is equal status among parties; 2) there is community support for the change; 3) common objectives are sought by both parties; 4) the connection is genuine and trusting, that is tutors should know their tutee well, and be interested in their success in the setting (Allport, 1954).

Peer tutor programs should be designed to reduce stereotypes and facilitate positive beliefs. Once these criteria listed directly above are met, research shows that knowledge about minority groups makes for more tolerant and friendly attitudes (De Boer, 2012; Slininger, 2000). Contact theory was conceptualized as a broad approach to affecting intergroup relations (Slininger, 2000) following the idea that knowledge about minority groups must be built through direct experience and credible information sources (e.g., sources that will provide accurate information).

Reasoning for the use of peer tutors and a peer tutor training program is that this construct fits nicely with the contact theory categories mentioned above. Slininger et al. (2000) suggest that there should be substantial planning in order to fit the program into the four categories of the contact theory construct. In the case of this study, there was considerable time spent discussing and planning with special education teachers in order to facilitate the most appropriate pairing for dyads, and the general education teachers in order to present the peer tutor concept in a positive and supportive manner as an opportunity to learn and help another student learn.

The purpose of this study was to evaluate the effects of a peer tutor training program and peer tutoring experience on the attitudes of peer tutors participating in a 15 session semester long peer tutor program. The research question the study sought to answer was whether or not peer tutor attitudes toward working with students with severe disabilities improved after peer tutor training and experience teaching the students with severe disabilities in an adapted physical education setting. The experimental design was mixed, including repeated measures and qualitative analysis.

Method

Twelve typically developing students from the same school were randomly selected to participate in the study from four classes; two 3rd-grade and two 6th-grade classes, 5 boys and 7 girls, age 9-12 (M = 10.33, SD = 1.44), were recruited to be peer tutors for this study. Because the peer tutors were released from academic learning sessions, the classroom teachers selected students to participate based on merit of academic performance, citizenship, and willingness to participate. If at any time the classroom teacher noticed a negative effect on peer tutor academic performance, they

were to be dismissed from the peer tutor program. No peer tutors were dismissed during this study. Peer tutors were assigned to dyads based on suggestions of the special education teacher. These dyads were static throughout the course of the study.

Three participants with disabilities were randomly selected from each of the four self-contained special education classrooms to take part in this study. Random selection was conducted by drawing names out of a container. Each of these participants had a current Individualized Education Plan (IEP) at the time of the study. The participants selected were, 7 boys and 5 girls, ages 6-11 ($M = 9.08$, $SD = 2.11$). The disabilities of the participants were intellectual disability, ASD, other health impairment, multiple disabilities, and/or developmental delay. All participants with disabilities were enrolled in a self-contained adapted physical education class for 30 minutes per week. The other students in those classes participated in the adapted physical education program but no data were collected on these students.

Instrumentation

Both quantitative and qualitative data collection protocols were followed. The dimensional components of attitude were assessed through the use of an attitude scale developed by the researcher. The scale was paired with verbal responses to specific cues. These instruments focus on the three components of attitude, cognitive, affective, and behavioral. Weekly verbal responses collected from each peer tutor participating in the study were used to determine the peer tutors' perceptions of their teaching experience, as well as their overall attitude toward their tutee for that day. This scale along with verbal responses was reviewed by experts who were both special education teachers, and university faculty. These experts agreed that these combined methods of measurement

were appropriate measures of attitudinal changes that may take place in the APE setting.

Visual Measurement of Attitude

The Likert-type scale with smiley faces to illustrate the points on the scale has been most commonly used with children in research environments (Macklin & Machleit, 1990). The "smiling face" scale was presented by Wells (1965) as a preference measure to be used with children 5-12 years old because they are more apt to use faces than words or numbers. Kunin (1955) reported the use of the smiley face scale as a viable option for efficient assessment of attitude toward a specific construct. The smiley face scale takes very little time to introduce because it is ubiquitous with elementary school settings (Sad, 2012). Given the time constraints, it was decided to pair the smiley face scale with a verbal response to a specific prompt at the end of each peer tutoring session.

The specific instructions for selecting a smiley face and describing the selection were "Please choose the smiley face that best reflects your attitude toward your tutee for today. Use the pen to circle one choice. Once you have circled a choice please speak into the microphone and state why you chose the smiley face you did." The process from selection to the end of verbal response averaged less than 2 minutes for peer tutors to complete.

To derive a score, each smiley image was assigned a numeric value: 1 = sad face, 2 = neutral face, 3 = happy face. Additionally a sad face is considered a negative attitude, a neutral face is neither negative nor positive, a smiley face is considered a positive attitude. Mean scores were calculated and compared for baseline and intervention sessions.

Verbal Responses

Open-ended verbal responses were also collected to have peer tutors offer further reasoning for their selection on the smiley face scale. After each session, peer tutors were asked to tell why they chose the specific smiley face on the scale mentioned above. These responses provided qualitative data from the peer tutors. These data were analyzed as being positive neutral or negative, to discern peer tutor attitude toward their tutee.

The basis for creating the Likert-type instrument and prompting peer tutors for verbal responses was to discern whether or not peer tutors felt they had a positive, neutral, or negative attitude toward their tutee. Prompting tutors to explain their answer serves to further explore their attitudes toward their tutee in APE.

Procedure

Data Collection

Data were collected in the APE teaching site from 12 peer tutors participating in four separate classes over the course of 15 sessions. Three participating peer tutors were in each class. All sessions were video recorded using a handheld iPad. All peer tutors wore the voice recorder and lavaliere microphone clipped to their shirt during each study session. For identification purposes, peer tutors stated the class, the date, and an assigned letter, which was a designation for their name, into the audio recorder before the data collection session began for identification purposes. Voice recorders were used to collect verbal interaction between peer tutor and tutee, as well as responses to the smiley face selection prompt. Audio data were coded from recordings after the sessions were completed for the day.

All study sessions took place in a multipurpose room. The room is the same one

in which the students currently participate in SPE. There was a schedule of activities for the day posted conspicuously on a wall for reference by the peer tutors, staff, and the teacher throughout the study.

Peer tutors who were selected to wear the microphones were prompted to meet with a graduate assistant at a table in a breezeway adjacent to the teaching environment after each session. Peer tutors were asked to make a choice on the smiley face instrument, and then were asked to tell why they made the specific choice they did. Efforts were made to maintain privacy during this process; however, there was not a protocol in place for absolute secrecy for peer tutor responses.

Baseline data were collected, and then peer tutors participated in a 30-minute peer tutor training session. Data were then collected for each session following intervention. Peer tutors also attended subsequent booster sessions prior to each teaching session. The Klavina et al. (2008) training protocol was adapted as a model for training peer tutors in this study. Peer tutors were trained in a single session due to significant demands on their academic schedule. These time constraints created significant issues for scheduling, but are a reality in public school education and therefore contribute to the generalizability of the study. The main training session focused on learning skills to be an effective peer tutor (e.g., appropriate communication, using tutee name, specific feedback, proximity, and system of least prompts). Peer tutors role played methods of teaching, providing feedback, and prompting tutees. Peer tutors completed a quiz at the end of the training session to evaluate recollection of concepts and to evaluate participant attitude toward the training session; however, a specific score was not required for participation in the study. Scores on the quiz ranged from 85% to 100%, with a mean of 92.5%. The training and

booster sessions were designed to offer peer tutors training in instructional assistance for tutees with severe disabilities in a clear and understandable manner. Topics of booster sessions were selected based on peer tutor evaluation scores from the previous week. Topics scoring the lowest were the focus of the booster session for the next week. Peer tutor objectives of communication, prompting, and feedback were conspicuously placed near teaching stations for reference during the session once intervention was implemented.

Design and Analyses

The data were collected using a multiple baseline design (Shadish, Cook & Campbell, 2002). This design allows the researcher to forego returning to baseline, which is useful when attitude is not expected to revert to baseline levels. Baseline data were collected for three sessions prior to introducing intervention with group 1. The intervention (peer tutor training) was introduced to group 1, while groups 2, 3, and 4 were still in baseline. Intervention was introduced to the remaining groups once its impact of was established for group 1 (O'Neill et al., 2011). Experimental control is established with successive introduction of the intervention across three different groups and points in time (Horner et al., 2005).

Visual analysis was used to inspect and interpret the data (O'Neill et al., 2011). Apparent changes in data patterns corresponding to the experimental manipulation of the intervention were determined by visual analysis. Noting the variability and trends within the conditions, as well as changes in level and overlap between conditions, is essential to determine the effect of the independent variable on the dependent variables (O'Neill et al., 2011). The less overlap between conditions, the smaller the variability within

conditions and/or the greater the level change. The greater the change in level, trend, and variability between conditions in the expected direction, the stronger the experimental control (Houston-Wilson et al., 1997). Immediacy of the effect of the intervention (e.g., how quickly the data change after introduction of the intervention) and the consistency of effects across legs of the baseline are important considerations for evaluating the effectiveness of the intervention in multiple baseline data (Horner et al., 2005).

This study focused on peer tutor attitude and whether or not it was affected by the peer tutor training and booster sessions. Baseline data and intervention data for verbal responses were compared using a Wilcoxson signed-rank test (Field, 2013). SPSS statistics package was used to analyze the quantitative data. The Wilcoxson signed-rank test is a nonparametric test used in this study because of the small sample size. In order to use this test, each statement was coded as positive, negative, or neutral. Once coding was complete, percentage scores for each category were calculated. There were deemed to be no neutral statements in the data so the category was omitted. The percentage scores for positive and negative comments between baseline and intervention were then analyzed using SPSS.

Qualitative data (the verbal responses from peer tutors after each tutoring session) were analyzed following the established procedures of (Bogden & Biklen, 1992; Drew, Hardman, & Hosp, 2007). Data were transcribed verbatim from the audio recordings into an excel spreadsheet, grouped by tutor, and analyzed to identify positive or negative terms and attitudes toward daily activities or tutees. Each comment was coded with a + or –. Fellow researchers audited the coding by reading the transcriptions and coding it with a + or -. Interobserver reliability was calculated as the number of agreements/agreements

plus disagreements, multiplied by 100. A predetermined mean of ≥ 80% interobserver reliability was deemed acceptable. Illuminating quotations were reported from each week's verbal responses to show peer tutor's thoughts and feelings about tutoring, working with a student with severe disabilities, and the classroom environment.

Data were collected in a multiple baseline across participants design (Kazdin, 1982) to evaluate peer tutor attitude as measured by the 3-point Likert scale instrument before (baseline) and after Peer Tutor Training (intervention). There were four separate classrooms from which tutors participated. Peer tutor selections for each group are presented in Figure 4.1.

Results

Group 1

Peer tutors in group 1 worked with tutees from a K-3 Functional Academics classroom. Tutees from this classroom were compliant and regularly participated in class without any behavior issues. The three tutees involved in the study were diagnosed with intellectual disability ($n = 1$) and other health impairment ($n = 2$). One tutor/tutee dyad from this group was removed from the study after week 12 due to the tutee moving out of state (Lisa, Rox). Data for this tutor were included, but were noted as incomplete.

Likert scale selections for Group 1 were collected in baseline for 3 baseline points. Mean levels across students were from ($M = 3.0$; $M = 2.3$), and standard deviation ranges from ($SD = 0$; $SD = 0.5$) without trend or variability (see Figure 4.1). Intervention phase data were collected for 10 sessions ($M = 3.0$; $SD = 0$). When comparing means, there is minimal change in level. There was little variability in the data, showing that Group 1 tutors made positive selections throughout data collection.

Group 2

Peer tutors in group 2 worked with tutees from a 4th-6th-grade Functional Academics classroom. The three tutees in group 3 were diagnosed with intellectual disability *(n* = 1), other health impairment *(n* = 1), and autism *(n* = 1). Two of the three tutees have rarely had any behavior difficulties in APE. The third tutee (Rex) exhibited some negative behaviors (not attending, not participating).

Likert scale selections for group 2 were collected in baseline for six baseline points. Mean ranges were from (*M* = 2.17; *M = 2.83*), and standard deviation ranges from (*SD* = 0.41; *SD* = 0.45) without trend or variability (see Figure 4.1). Intervention phase data were collected for 10 sessions (*M* = 3.00; *SD* = 0). When comparing means, there is no change in level. There was no variability in the data, showing that Group 2 made positive selections throughout data collection.

Group 3

Peer tutors in group 3 worked with tutees from a K-3 functional life skills classroom. Tutees in this class presented unique challenges for the peer tutors. The 3 tutees involved in the study were diagnosed with developmental delay. Each tutee had a behavior management plan to assist tutees in transitioning, following directions, and participating, APE. There were at least five adults present to support peer tutors and tutees during the study.

Likert scale selections for group 3 were collected in baseline for 9 baseline points. Mean ranges were from (*M* = 2.70; *M = 2.89*), and standard deviation ranges from (*SD* = 0.38; *SD* = 0.48) without trend and minimal variability. Intervention phase data were collected for 5 sessions (*M* = 3.0; *SD* = 0). When comparing means, there is no change

in level. There was no variability in the data, showing that Group 3 made positive selections throughout data collection.

Group 4

Peer tutors in group 4 worked with tutees from a 4-6 functional life skills classroom. The three tutees involved in the study were diagnosed with other health impairment ($n = 1$), intellectual delay ($n = 2$). These three tutees sometimes struggled with transitioning, following directions, and attending in APE. Each tutee has a behavior management plan created by the classroom teacher. During this study there were at least four adults present at each class to support tutors and tutees.

Likert scale selections for group 4 were collected for 12 baseline points. Mean ranges were from ($M = 2.82$; $M = 2.91$), and standard deviation ranges from ($SD = 0.30$; $SD = 0.40$) without trend and minimal variability. Intervention phase data were collected for 5 sessions ($M = 3.0$; $SD = 0$). When comparing means there is no change in level. There was no variability in the data, showing that group 4 made positive selections throughout data collection.

Evaluating peer tutor attitude selections across all 4 groups presents an opportunity to consider replication. Groups showed a decrease in variability from some during baseline to none during intervention, with the exception of group 1, where Sam and April chose neutral rather than positive for their attitude selection in some sessions during intervention. The results across groups do not support an effect for the intervention due to a lack of level change, and complete overlap between phases due to a ceiling effect. These results do show that peer tutors made positive selections in baseline and intervention phases, suggesting that the peer tutor training did not have an adverse

effect on peer tutor attitude.

Verbal Response

Findings are reported for the differences between peer tutor attitude in baseline and intervention phases based on verbal responses. Table 4.2 shows the percentage of positive and negative comments from peer tutors for baseline and intervention phases.

The results of the Wilcoxson signed-rank statistical analysis showed that peer tutors made significantly more, positive comments in intervention (M*dn* = 82.58) compared to baseline (M*dn* = 68.17), $t = -1.96$, $p = .050$, $r = .28$.

Qualitative Findings

One purpose of asking students to offer verbal responses was to obtain feedback concerning their smiley face selection and what they liked and disliked. The rationale for collecting these data was to add a level of explanation for the smiley face selection the peer tutor made. Table 4.1 shows the mean and standard deviation for the Likert scale selections. We thought that two evaluations would provide more detail for evaluating peer tutor attitudes.

The specific instructions given to students were "Please say why you made the smiley face choice you did." Following is a list of illuminating statements peer tutors made about their tutees throughout the study.

Week 1. "I love Linn she's my best. Give this (pedometer) to my buddy? What if my buddy is in a wheelchair? Hi Linn, what's up sis?"

Week 2. "Today went great, Rex stayed at all his stations and he wasn't wandering around."

Week 3. “Rox is a very very good peer tutee, she is really special and I like throwing stuff with her.”

Week 4. "Eddy tackled Susie today. I didn't know what to do. A teacher came over and told him to have nice hands. I circled a flat face"

Week 5. “Today was a little better, Rex stayed at the first station the whole time – that has been our goal, so I am choosing a smiley face.”

Week 6. “It was good – it was fun and Mel had a great time – I think.”

Week 7. “I think Gus did a good job today. I did a smiley face because I thought Gus was listening – he was listening OK – he kept throwing his running thing [pedometer], he still was participating good so, I am choosing a smiley face.”

Week 8. “Today went well. Today was the best I think Rex had ever done.”

Week 9. “I think today went really well. Linn was really interactive.”

Week 10. “I’m gonna do a smiley face, because Louise is the best partner I could ever ask for. She is very very kind.”

Week 11. “I really enjoyed Gus today because he followed directions and he had a really good time and I liked playing with him today.”

Week 12. “I think Avery did OK today, and she actually conversated [sic]. Avery is awesome.”

Week 13. “Today I am circling a happy face because, just for some reason, I feel greater [sic]. Pat talked. So I am circling a smiley face. I just had a great day.”

Week 14. “I had a great time with Eddy. He was really fun. He loved the scooters. Overall he was just awesome.”

Week 15. “Glad that Susie is feeling better than last time. It was really fun to play

with them and it seemed like they were happy."

Week 16. "I had a great time with Mel. This is probably the first time getting him on the scooter."

Discussion

The purpose of this study was to evaluate the effects of a peer tutor training on peer tutor attitude. These findings support the importance of a training program for peer tutors in APE in order maintain and improve the attitudes of peer tutors toward peers with severe disabilities.

Of the eight threats to internal validity of an experiment (Thomas et al., 2015), history (events occurring during the experiment that are not part of the treatment) was the greatest challenge. Tutors had contact with tutees in the hall, at recess, lunch, and specialty classes (e.g., music, library, art, etc.). Moreover, tutors had contact with one another outside the APE class. These contacts could not be controlled. Thus, many variables, inside and outside the school, may have contributed to improvement in tutor attitudes. The significant Wilcoxson signed rank test showed an increase in positive comments when comparing rates between baseline and intervention for both males and females inclusive of all ages. These results support the idea that peer tutor training contributed to the attitude change, but should be interpreted with caution as they were not experimentally controlled. These results could have also been confounded by maturation or social desirability of responding positively.

Peer tutor training was supported by the significant positive results in the study. These results were surprising because the training and booster sessions were very limited in duration. In retrospect, however, the inability of this researcher to control extraneous

factors in a real-life setting may explain why tutor attitude results were so positive. Additionally, special educators at this school diligently work to create opportunities for students with disabilities to be included in daily activities with their same age peers and this could have led to the increase in attitudes.

Contact theory is a sound construct from which to assess a peer tutor program and peer tutor interactions with students with severe disabilities in APE. However, many pedagogical decisions were made empirically because there was little research to guide these decisions. For example, peer tutors were assigned tutees based on feedback from special education classroom teachers. There is a question of whether or not peer tutors and/or tutees should have been able to self-select, hence adding more autonomy into the tutoring environment. Autonomy is an important aspect of contact theory (Allport, 1954) because it allows for more choice and freedom in the setting.

A recommendation to strengthen the qualitative aspects of future research would be to reevaluate procedures for verbal data collection. The use of verbal responses to add depth to the peer tutor smiley selection was an important aspect of the study. These responses may have been more successful if peer tutors had a more private setting to record their responses. These responses have a "group think" (Drew et. al., 2008) quality due to their similarities. Moreover, the 3-point Likert instrument consisted of only three choices. There should have been at least 5 choices for students to select. More choices would have added more nuance to the data to be able to distinguish more between selections.

Analysis of the statements made about students with severe disabilities over the duration of the study revealed that many of the interactions were pleasant, rewarding, or

meaningful. These responses generally became more detailed as the study continued. It is clear that the peer tutors began to see such qualities as being fun, funny, or happy in the students with severe disabilities as the study progressed. These descriptive words signify positive attitudes in the peer tutors. This provides support for the fact that peer tutors related the tutoring experience as positive, even though it was difficult at times. It is unknown how a study of longer duration, perhaps 9 months instead of 4 months, would have affected peer tutor attitudes toward their tutees.

Conclusions

Based on the majority of the quantitative and qualitative findings of this study, it was concluded that peer tutor attitudes may have been positively affected by the peer tutor training intervention. These peer tutors all shared varying levels of improvement in attitude toward students with severe disabilities. A semester (4.5 months, 16 study sessions) of weekly contact in skill building and cooperative games provided enough time for peer tutor attitudes to improve. The findings of the present study do not support the contact theory of Allport (1954), who suggested contact would change attitudes in a positive direction only when the contacts were equal status. In the case of this study and others (Ensergueix et al., 2010; Tripp, French, & Sherrill, 1995), peer tutors were given additional status and responsibility for tutoring their tutee. On the other hand, the relationships were cooperative, intimate, and supported with community sanction. Nevertheless, peer tutor attitudes were positively affected by the experience with students with severe disabilities in APE.

Table 4.1: Peer tutor 3-point likert score selections: mean and standard deviation.

Group	Baseline		Intervention	
	M	*SD*	*M*	*SD*
Group 1				
Sam	2.33	0.58	2.50	0.52
April	2.33	0.58	2.92	0.29
Lisa	3.00	0.00	2.92	0.29
Group 2				
Dan	2.83	0.41	3.00	0.00
Pre	2.80	0.45	3.00	0.00
Wayne	2.17	0.41	3.00	0.00
Group 3				
Evan	2.89	0.33	3.00	0.00
Dax	2.80	0.42	3.00	0.00
Misty	2.70	0.48	3.00	0.00
Group 4				
Mel	2.91	0.30	3.00	0.00
Abby	2.82	0.40	3.00	0.00
Tim	2.91	0.30	3.00	0.00

Table 4.2: Percentage positive and negative peer tutor attitude – baseline, intervention.

Group	Baseline		Intervention	
Group 1	% Positive	% Negative	% Positive	% Negative
Sam	0	100	55	45
April	100	0	100	0
Lisa	66	34	72	28
Group 2				
Dan	100	0	100	0
Pre	50	50	50	50
Wayne	100	0	100	0
Group 3				
Mel	54	46	75	25
Tim	64	36	75	25
Abby	64	36	100	0
Group 4				
Evan	66	34	100	0
Dax	88	12	66	33
Misty	66	34	100	0

Note: Percentage scores out of 100% possible

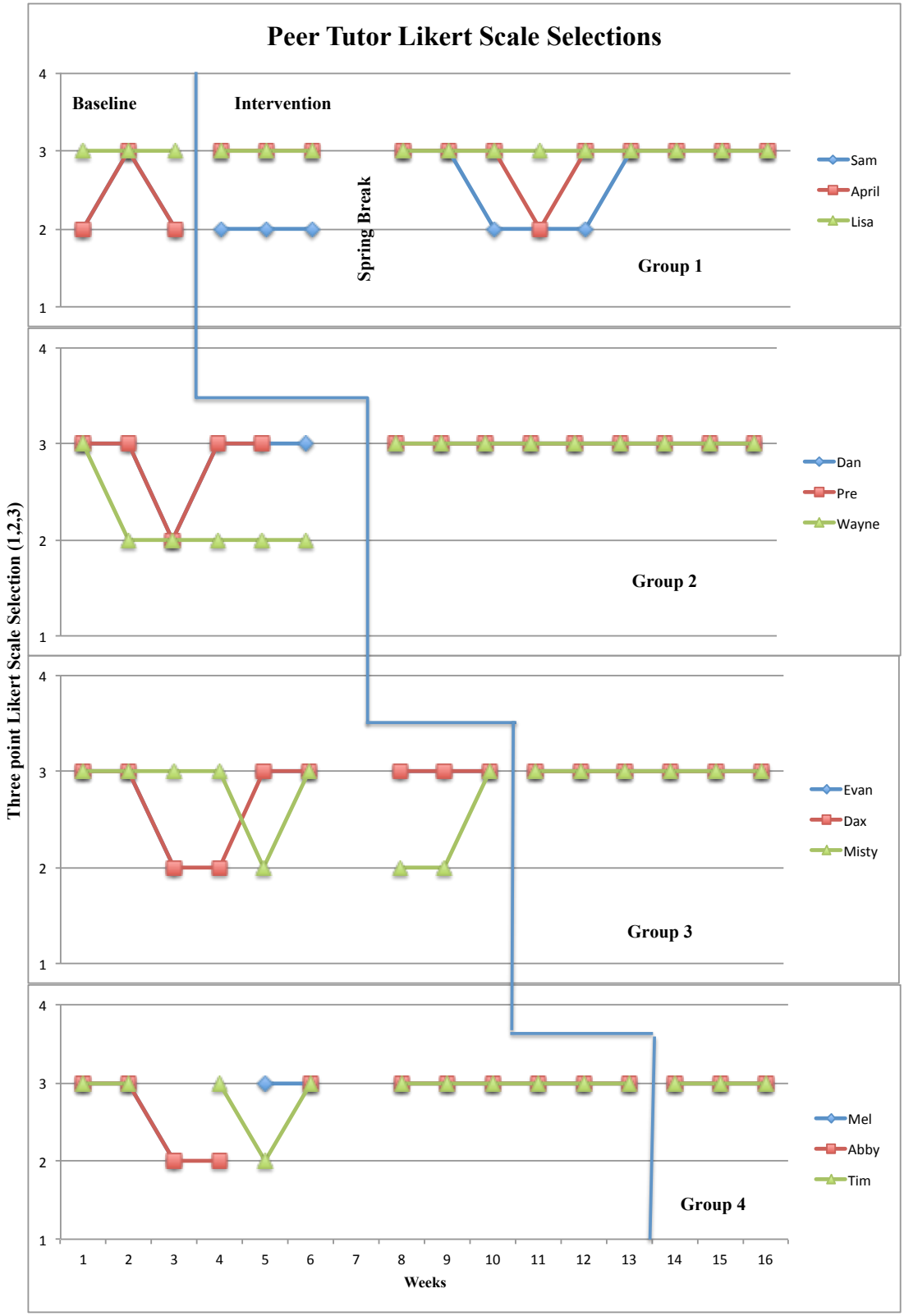

Figure 4.1: Peer tutor 3-point Likert scale attitude selections.

CHAPTER 5

CONCLUSION

Research into peer tutor programs for adapted physical education settings is a worthwhile endeavor. There are significant amounts of time that students with severe disabilities spend in a cohort with little interaction with students without disabilities. Likewise, students without disabilities often do not have opportunities to interact with students with severe disabilities. This study sought to illuminate the necessity for a peer tutor training protocol in order to make these interactions more effective and meaningful for both peer tutor and tutee.

Major Findings

The studies included in this dissertation support several major findings: (1) the initial development and content validity evaluation of the Peer Tutor Evaluation Instrument (PTEI), (2) the development of a peer tutor training protocol, (3) the determination that the peer tutor training protocol can be an effective way to improve teaching behaviors of peer tutors as measured by the PTEI, (4) the determination that trained peer tutors can have a positive effect on tutee performance of motor skills as measured by the TGMD-2, (5) the determination that the trained peer tutors can have an effect on activity levels of tutees, (6) the determination that peer tutor attitudes can be improved by peer tutor training. These six major findings support practitioner use of peer tutor training in adapted physical education (APE).

Finding 1

This study reported the development and initial content validity outcomes for the PTEI, a peer tutor performance evaluation in adapted physical education settings. Ratings of external expert reviewers provided essential feedback for the development and

refinement of the PTEI.

External experts provided content-related evidence of validity for the PTEI. A majority of the experts rated each PTEI item as relevant and concurring with the defined categories (communication, prompting, feedback).

Finding 2

It was determined that the peer tutor training protocol (PTTP) can be an effective way to improve teaching behaviors of peer tutors as measured by the PTEI. Performance scores for peer tutors improved when baseline and intervention phase scores were compared.

Finding 3

The relative improvement in performance of all peer tutors after intervention supports the prediction that peer tutor training is an important aspect of a successful peer tutor program in APE. These improvements compared favorably with previous studies, (Klavina, 2001; Houston-Wilson, 1997) showing that elementary age peer tutors can successfully teach skills to students with disabilities in self-contained SPE classes, or an inclusive PE class. A substantial body of literature supports using peer tutors to teach students in SPE/APE (Barfield et al., 1998; Block, 2007, 1994; Block et al., 1995) with a focus on improving motor skills. This study also supports the need for tutors to engage in training (d'Arripe-Longueville et al., 2002; Houston-Wilson, 1997; Klavina, 2001) to be as successful as possible engaging with and teaching students with severe disabilities in SPE. Other studies have proposed longer training sessions for peer tutors working with students with severe disabilities (Klavina, 2002; Long et al., 1980). This study supports

findings from Houston-Wilson (1997) by showing that elementary-aged students can be taught to tutor students in a short amount of time. This realization is important because of the increasing time constraints of students in elementary school (e.g., standardized testing, specific academic support, field trips, educational fairs, etc.). These findings show that if a small amount of time can be set-aside for training accompanied by booster sessions a few minutes before each class, there is a great potential for improvement of peer tutor teaching performance.

Finding 4

Students with severe disabilities in this study demonstrated delayed motor skills relative to their typically developing peers. The students demonstrated a mean TGMD-2 score of 22% for locomotor skills and 28% for object control skills. Generally, these scores mean students can perform one out of four criteria for demonstrating a motor or object control skill. Peer tutoring had a statistically significant effect on tutee performance of motor skills (i.e., skip, gallop); however, no such effect was found for object control skills (i.e., overhand throw, underhand roll).

Finding 5

Tutee activity levels (number of steps) as measured by pedometers were affected by the intervention. Multiple baseline data for baseline and intervention showed little experimental control, as well as high rates of overlap and variability. The results do show that tutee activity levels are consistently high throughout the study, and there is no sacrifice for activity with implementation of a peer tutor program.

Finding 6

Based on the quantitative and qualitative findings of this study, there is preliminary evidence that peer tutor attitudes are affected by a peer tutor training intervention. These tutors all shared varying levels of improvement in attitude toward students with severe disabilities. A semester (4.5 months, 16 study sessions) of weekly contact in skill building and cooperative games provided enough time for peer tutor attitudes to improve. The findings of the present study do not support the contact theory of Allport (1954), who suggested contact would change attitudes in a positive direction only when the contacts were equal status. In the case of this study and others (Ensergueix et al., 2010; Tripp, French, & Sherrill, 1995), peer tutors were given additional status and responsibility for tutoring their tutee. On the other hand, the relationships were cooperative, intimate, and supported with community sanction. Nevertheless peer tutor attitudes were positively affected by the experience with students with severe disabilities in APE.

Analysis of the statements made about students with severe disabilities over the duration of the study revealed that many of the interactions were pleasant, rewarding, or meaningful. These responses generally became more detailed as the study continued. It is clear that tutors began to see qualities (e.g., fun, funny, happy) in students with severe disabilities as the study progressed. These descriptive words signify positive attitudes in the peer tutors. Equally important, 90% of the peer tutors rated the tutoring experience as positive. This provides support for the fact that peer tutors related the tutoring experience as positive, even though it was difficult at times. It is unknown how a study of longer duration study (9 months in lieu of 4 months) would have affected peer tutor attitudes

toward their tutees.

Limitations

Finding large numbers of participants in special education research is often a difficult task. In the case of this study, 12 students with disabilities were recruited, as well as 12 students from general education. It is difficult to generalize results due to the small number of participants in the study. Moreover, effect size is often very small. Finally, this study was conducted in self-contained classroom settings where participants can be unpredictable, leading to high amounts of variability in the data. Some of this variability may have been the result of instrument limitations.

One of the limitations of the PTEI is the fact that there are only nine objectives for evaluating the myriad behaviors and interactions that occur between tutor and tutee in the APE setting. The addition of another objective would create a larger cross-section of peer tutor prompting for consideration, adding to the strength of the content validity for PTEI. This instrument is designed to provide a view of peer tutor performance, but does not encompass the entirety of variables present in the situation.

Discussion

The results of this study serve to provide an important initial probe connecting theoretical and pragmatic considerations to practical issues regarding the quality of peer tutor performance in APE. The PTEI marks a beginning to the process of providing a quantitative measure to evaluate peer tutor performance in APE. In summary, there is not one all-encompassing instrument to observe and record all aspects of peer tutor performance during APE. A multitude of external factors can influence peer tutor

performance (e.g., age of students, severity of disability, experience of APE teacher, opportunities for interaction, and structure of the APE curriculum). The PTEI presents an addition to other observation systems analyzing peer tutor performance in APE learning environments.

The peer tutor training program and PTEI assessment were developed for this study and were implemented to create an opportunity to train and evaluate peer tutors. The peer tutor training program protocol focused on teaching peer tutors how to establish a rapport with students (e.g., talking, interacting, proximity). It also focused on prompting and specific feedback related to tutee performance and behavior. For example, prior to training, tutors would simply say "good job," or not say anything to their tutee for the entirety of the class. After training, tutors would give specific performance feedback such as, "Sven I like how you kicked the ball at the target. This time try to kick it with the inside of your foot." At this point, the tutor was touching the inside of the tutees foot and showing them where to contact the ball. When training peer tutors, it is important to use physical prompts when necessary to help tutees both hear about and see the correct elements of a skill. Specific feedback from peer tutors seems to be more meaningful coming from a peer tutor who is trained to establish rapport with their tutee. This type of attention is difficult for classroom teachers to give because of diverse needs from students, large class sizes, and significant time constraints. Thus, trained peer tutors can offer much needed support to a learning environment for students with severe disabilities in APE.

Recommendations

Future studies should examine the effect of a more lengthy peer tutor training (3-4 sessions) to improve peer tutor performance. Other variables that might be measured in peer tutors following training would be whether peer tutors feel more confident, less nervous, happier, or more empathetic while working with their tutee. Further research could also consider another school setting with different age groups, such as junior high or high school, to begin examining how to best train tutors of different ages. Finally, similar studies should be repeated over a longer period of time to determine if peer tutor's teaching skills continue to improve over time and if tutees continue to improve their motor skills and number of steps taken during the session.

Recommendation to strengthen the qualitative aspects of future research would be to reevaluate procedures for verbal data collection. The use of verbal responses to add depth to the peer tutor smiley selection was an important aspect of the study. These responses may have been more successful if peer tutors had a more private setting to record their responses. These responses have a "group think" (Drew et. al., 2008) quality due to their similarities. Moreover, the 3-point Likert scale instrument consisted of only 3 choices. There could have been at least 5 choices for students to select. More choices could have added a depth to the data that may have explained more attitude change.

APPENDIX A

SCORING PROTOCOL – PEER TUTOR EVALUATION INSTRUMENT

Scoring Protocol: Peer tutor evaluation instrument

Instrument Overview

The instrument is designed to evaluate essential skills for peer tutors in SPE. It is not comprehensive and may not consider some skills students exhibit in the setting.

Scoring Procedures

The evaluation instrument should be used while directly observing live class or video recordings of peer tutors working with their tutees (video recordings may be preferred in order to evaluate multiple peer tutors in a single session). There is not a designated time limit for the observation session, but there are some objectives that may not be met until the session is completed.

Verbal interaction variables can be difficult to evaluate in a loud physical education space. Wireless voice recorders can help collect these data. If they are not available, observers need to be close enough to listen to tutor / tutee interactions.

Scoring

Many objectives are evaluated on a 4-point scale (1=Needs significant improvement, 2=Emerging skill, 3=Skill is demonstrated in some situations, 4=Skill is demonstrated appropriately over all situations) with a total possible score of 28. If no physical prompt is needed, the total possible score would be 24.

Scoring Individual Objectives

When scoring objectives, the observer should consider each objective individually and be aware of the context the students are in.

Goal 1: Appropriate engagement and interaction with the peer tutee

Objective 1: Peer tutor will stay within 5 feet of their tutee during the entirety of the skill session

Peer tutor should be scored on the amount of time he or she spends with the tutee. Peer tutors should stay within approximately 5 feet (normal voice talking distance) in order to maintain engagement.

It is important to note that the observer must use some discretion when assessing this objective. If the peer tutor leaves their tutee to run and grab an errant ball and comes directly back to practice again, the score should not be negatively affected. Tutors will be scored a 0, 1 based on the amount of engagement with the tutee. If the tutor turns from the tutee or leaves to engage in an alternative activity for more than 10 seconds, they will be scored a 0. If the tutor maintains engagement for the entire session, they will be scored a 1.

Objective 2: Peer tutor will ask at least two general check-in questions during the session

Peer tutor should be scored on whether or not they ask two general check-in questions of their tutee (e.g., "How are you today Johnny?", "What did you have for lunch today Johnny?"). Tutors will be scored (0,1,2) based on the number of questions asked.

Goal 2: Appropriate use of peer tutee name

Objective 1: Peer tutor will use peer tutee name at least four times during the session

The number of times the tutor uses the tutee name should be counted up to 4. Peer tutor will be scored (0,1,2,3,4) based on the number of times the tutee name is used.

Goal 3: Appropriate verbal direction for tutee to perform task and provide processing time

Objective 1: Peer tutor will use correct verbal prompt / demonstration to begin practice session

If the tutor provides a correct prompt, (e.g., "Johnny it is time to kick the ball.") and demonstration of the skill (mature pattern for skill) to begin the practice task at least once, a score of 1 is assigned. If they do not exhibit the appropriate cue / demonstration they are scored a 0.

Objective 2: Peer tutor will wait at least 10 seconds before giving another direction or asking a question

The second objective monitors processing time.

Peer tutors need to wait at least 10 seconds before giving another prompt. If tutor waits at least 10 seconds, a score of 2 is assigned. If they do not, a score of 0 is entered.

Goal 4: Appropriate use of a physical prompt when necessary

Objective 1: Peer tutor will use a physical prompt only after a verbal prompt and a demonstration have been given

The score for this objective (N/A,1) varies according to set criteria. If the peer tutee performs the task correctly and a physical prompt is not required, a score of 'N/A' is entered. If a physical prompt is required, the observer should first note that the peer tutor gave a verbal prompt and demonstration; in this case, a score of 1 would be earned.

Goal 5: Appropriate positive specific performance feedback

Objective 1: Peer tutor will offer at least four instances of positive specific performance feedback during the session

The peer tutor should offer at least four separate instances of specific positive performance feedback (e.g., "Nice kicking the ball with the instep of your foot, Johnny." "Johnny, you are doing great catching the ball with your hands.") while the tutee is practicing the skill. Scoring (0,1,2,3,4) depends on the number of separate instances of feedback provided.

Goal 6: Appropriate positive specific behavior feedback

Objective 1: Peer tutor will offer at least four different instances of positive specific behavioral feedback during the session

Four instances of specific positive behavior feedback should occur in the course of the class (e.g., "Johnny, you are doing great sitting and listening to the teacher." "Johnny, you've had nice hands for the entire class!"). Observers should note instances of positive specific behavior feedback up to 4. Scoring (0,1,2,3,4) depends on the number of separate instances of feedback provided.

Goal 7: Social interaction

Objective 1: Peer tutor will communicate either verbally or nonverbally with the tutee on content not related to instruction (i.e., comments on dress, comments to engage, compliments, high fives, etc.) at least four times during the session

The observer should note instances in which the peer tutor engages in social interaction behaviors with the peer tutee. These behaviors can include verbal or nonverbal communication such as, comments on dress, comments to engage, compliments, or high fives.

APPENDIX B

PEER TUTOR TRAINING PROTOCOL

Peer Tutor Training Protocol (PTTP)

Description of Main Peer Tutor Training Session

The purpose of the main peer tutor training session is to teach peer tutors to interact with students with disabilities using specific steps (Klavina, 2010). These steps are intended to promote communication, prompting, feedback, and social interactions between students with disabilities and peers without disabilities in special physical education. A secondary focus of the peer tutor training is to help students become more comfortable helping students with disabilities. The peer tutoring steps form a set of interactive strategies to: (1) *communicate* – provide teaching instructions to their tutee; (2) *prompt* – provide verbal and physical assistance to their tutee; (3) *feedback* – offer specific suggestions for improving performance and behaviors; (4) *play* – participate in physical education activities together with their tutee (Klavina, 2010). The researcher will conduct one 30-minute tutor training session. This tutor training session will be supported by 3-5 minute booster sessions (described at the end of the main session) prior to each class tutors participate in. The 30-minute training session is designed to provide an overview of the three main domains tutors will be evaluated on using the Peer Tutor Evaluation Instrument (PTEI) when they tutor (Communication, Prompting, and Feedback).

Materials required for training: Paper, writing utensil, ball for each group (approximately 5-7), instructor script, open space for tutors to role play.

Main Session (30 Minutes)

In the main session, the instructor should provide general information about peer tutoring for students with disabilities. Next, the instructor will teach three steps: (1) communication (engagement and interaction with peer tutee, use of peer tutee name), (2) prompting (verbal directions, and SPE prompt), and (3) feedback (positive specific performance, positive specific behavioral, and social interaction). The practitioner should be cognizant of time constraints due to the amount of content that is suggested to be addressed in this session. It is suggested that the number of participants in the training session be limited to 15 to reduce potential need to use time to manage behaviors, or organizational needs.

General Information: Empathy

The main session also focuses on how people are unique in order to help students learn empathy.

Instructor Script: "What is a tutor? A tutor is similar to a teacher; because they are good at something and can help others learn how to do it. Peer tutors are teaching something new to a friend. For example, when your friend teaches you how to kick a ball, you are the tutee – because you are learning from someone else. On the other hand, you are the tutor when you teach your friend to jump higher. This means that all of us are learning and teaching without even thinking about it. The reason we can be tutors or tutees is because we are all teachers and learners. The amazing thing about each of us is that we all have unique skills that we can teach one another. Did you know that you can learn from your tutee as well? "

Peer Tutor Protocol

Three main categories will be discussed (communicating, prompting, and feedback). Appropriate cueing (e.g. "We are going to kick the ball."), modeling (e.g., the tutor demonstrates how to kick the ball) and providing specific feedback will also be presented.

Instructor Script

Communicating: "The first step to being a good peer tutor is to know how to talk to your tutee. Be sure to use a clear voice and remember that you should make sure your tutee is listening to you. Ways to check are; if your tutee looks at you, follows the direction you gave, gives you a high five or 'bones', etc. There are three goals for great communication with our peer tutee.

Appropriate Engagement and Interaction With the Peer Tutee

You need to try and stay within 5 feet of your tutee during the skill session. Do you know how far 5 feet is? It is a little taller than you are. If you are close to your tutee there is a better chance they will hear you and respond to you. It is OK to run and grab a ball or other object that you are playing with, as long as you come right back to play again. It is important to give your tutee some space if you think they need a break from the activity.

You need to ask at least two check-in questions at the beginning of the class using your peer tutee's name. You could say something like, "Hey Johnny how are you today?", or "Johnny, what did you have for lunch today?". If your peer tutee does not talk you can just say, "It is great to see you today Johnny!", or "Johnny, we are going to work on kicking today!").

Appropriate Use of Peer Tutee Name

"It is also important to use the name of your peer tutee when you are talking with them. You should try to use their name each time you ask them to do something or want to talk with them. A great goal is to use their name at least 5 times in the time you are working with them. Chances are you will use their name many more times than that!"

Role Play Practice

"Now we are going to take three minutes to practice what we just talked about. Please group in pairs of two and practice some creative check-in questions and statements. Use the paper I gave you to write your favorite down to share with the group."

Prompting: "Prompting means any verbal or physical instructions from you (peer tutor) to the tutee related to practicing in the SPE setting. What are some examples of a verbal prompt? What about a nonverbal prompt? What about a physical prompt?"

Appropriate Verbal Direction for Tutee to Perform Task and Provide Processing Time

"It is important that your tutee know what they need to be doing during the SPE class. You should tell your tutee in clear and simple terms what they need to do (e.g., "Johnny you need to stand up" or "Johnny we are going to practice kicking. You need to kick the ball to me"). The second step with prompting is that you need to wait 10 seconds (you can even count in your head) before asking your tutee to do the something else – or the same thing again. Now we should talk about a physical prompt in special physical education."

Appropriate Use of SPE Physical Prompt When Necessary

"A physical prompt in this class means touching your tutee with a "hand-over-hand" prompt so you can help them feel the correct movement associated with throwing a ball overhand, kicking the ball with the instep of their foot, spreading their fingers to catch a ball, etc. How could you physically prompt your tutee to kick a ball with the inside of their foot? One important thing about physical prompts is that you can only use one if you have told and shown your tutee what to do first. So remember – first you tell your tutee what to do, then you show them what to do – after this you can use a physical prompt."

"Remember one last thing about a physical prompt; you should not be physically prompting tutees who are misbehaving (e.g., picking your tutee up off the floor because they are refusing to participate, etc.). In this case, you should contact the teacher or staff for help."

There are three goals for excellent prompting with your peer tutee.

Objectives

1. You should use the correct verbal prompt to begin the skill practice session (e.g., "Johnny, it is time to practice kicking. You need to kick the ball"). Make sure you use clear and simple words to talk to your tutee.

2. You need to be patient and wait for your tutee to do what you have asked them to do. Count to 10 in your head before telling them what to do.

3. You should use a physical prompt only when necessary. If you peer tutee does not perform the task correctly after the use of a verbal prompt and demonstration, you can show them using a physical prompt. Remember that if your tutee refuses to follow your direction, you can ask a teacher for help.

Role Play Practice

"Now we are going to take 3 minutes and practice the correct way to offer prompts to begin practicing. Please make groups of two and practice being the tutor and tutee. Make sure you use clear and simple words. – Now practice correct ways to use a physical prompt. Remember you first need to tell, and then show, finally you can use a physical prompt. Can we have a group come to the front and show us the correct way to prompt a peer tutee?"

Feedback

The last topic we will talk about today is how you help your tutee know they are doing something right in the class. Another way to say this is 'feedback.' You get feedback from friends, parents, and teachers all the time. Can anyone give an example of feedback? What about positive feedback? What about negative feedback? It is important that your tutee know when they are doing something right, and you need to make sure you tell them. There is a special way we are going to learn to tell our tutee they are doing something right. You need to say, "Johnny, great kicking the ball with your foot!" instead of "Great job Johnny!" What is the difference between these two? The first one tells your tutor what they are doing great at. The second does not. Another example is "Johnny, you are doing great listening to the teacher!" Remember positive specific feedback for your tutee. There are two goals for great feedback for your tutee.

Objectives

1. You need to try and tell your tutee at least four times that they are doing great. Remember to make it specific and positive (e.g., "Johnny, great kick to the target with the

instep of your foot." or "Johnny, you kicked the ball with your toe. Now you need to practice kicking with the instep of your foot.") (Klavina, 2008).

2. You need to try and tell your tutee at least four times that they are behaving great in the classroom (e.g., "Johnny, I love how you are practicing kicking the ball!", "Johnny, thanks for listening to me today!" etc.).

Role Play practice

Now we are going to take 5 minutes and brainstorm some creative ways to offer positive feedback to our tutee. Please make groups of two and practice being the tutor and tutee. You are now going to practice a specific skill (tossing a bean bag). Tutors need to provide at least 4 different positive feedback statements to their tutee.

End of Main Practice Session

Peer Tutor Training Booster Sessions

Booster sessions should be designed to address areas the practitioner identifies that peer tutors should be improving on when working with their tutee. These sessions should be short (2-3 minutes), clear, and directly relate to the objectives in the PTEI. One method for presenting content to tutors is through the use of examples and nonexamples. Each booster session and instructor script for PTEI objectives is enumerated below.

Communication

Appropriate Engagement: "Remember you need to stay close to your tutee when you are teaching them. Who remembers how close you need to stay to your tutee? Is this a good distance? (Instructor should be a significantly longer distance than 5 feet away from the group of tutors). Tutors should answer 'no'. (Instructor should now show students the suggested five-foot distance.) This is the correct distance for you to think about when teaching your tutee. Remember that it is OK to go and get a ball, or any other object you are working with as long as you come right back and continue to work with your tutee."

Use of Tutee Name: "Who remembers when to use your tutee's name? Do we repeat their name over and over? "Johnny, Johnny, Johnny, Johnny, Johny." Tutors should answer 'no'. Or do we say the tutee's name clearly each time we need them to do something? This is a good example, "Hey Johnny, let's practice kicking the ball!" Remember to use your tutee's name at least 5 times."

Prompting

Use of Appropriate Task Prompt and Wait Time: "Remember that you need to tell your tutee what they need to do in a clear and friendly way. Should we ask them if they want to kick the ball? What happens if your tutor says 'NO'? Remember to use a statement such as "Johnny, let's go learn about throwing and catching at Station 1." This way you are giving a clear and understandable instruction to your tutee. Remember to ask for help from a teacher if your tutee does not follow the direction you gave her. Is it OK to just keep asking your tutee to do something over and over again? "Johnny you need to kick the ball, Johnny you need to kick the ball, Johnny you need to kick the ball." Remember that your tutee can get overwhelmed with instructions and you need to be patient and wait at least 10 seconds before giving additional instructions. You can even count to 10 in your head to make sure you are waiting long enough."

Use of Appropriate SPE Physical Prompt: "We need to remember that it is OK to help our tutees feel the correct way to perform a skill. Can I just grab the tutee and move their hand to throw overhand? That may make you mad and could do the same for your tutee. Remember you need to ask them to try throwing, show them how to throw, and if they are not performing the skill correctly, tell them that you are going to move their arm so they know how it feels to position it to throw the ball. (At this point, the instructor should be physically touching a tutor's arm to show tutors how to contact tutees). Be gentle, and never make a tutee do something they do not want to do."

Feedback

Specific Performance Feedback: "Remember that, if your tutee needs to improve a part of their practice, you need to tell, or show, them **specifically** what they should

improve. Can you just say "Good Job?" Tutors should say 'NO'. For example, if you are kicking a ball back and forth with your friend and he/she is kicking the ball with their toe, you could say, "Johnny, I like how you are stepping with your foot, now let's see if you can kick the ball with the **inside** of your foot.", or "Johnny, you are doing awesome following directions, now try kicking the ball with the **inside** of your foot."

Specific Behavioral Feedback: "It is important to tell your tutee that they are behaving well in the class. Be sure to **specifically** tell them what they are doing well in the class, "Johnny, it is great that you are sitting and listening to the teacher." Remember that your tutee may behave inappropriately. It is important to tell them what they need to do differently ("You need to listen." or "Your hands are hurting me– you need to stop.") are two examples you can use. There can be different reasons your tutee is not behaving appropriately. They may not understand the directions, may not want to follow the directions, or may want attention from you. Sometimes it is best to tell your tutee what they need to do and then wait for them to do it. Do not give them attention. Be patient and wait for them to follow the directions. If you ***ever*** feel uncomfortable with what your friend is doing, please tell a teacher and we will help you."

Social Interaction Behavior: "It is important for you to get to know your tutee. They may want to follow directions and work with you if they know you like working with them. Try to make friends with your tutee. Ask them questions about what they like to do, "Do you like to play baseball?", "What is your favorite sport?" You can also tell them something nice, "I really like your shoes.", "Did you get a new haircut? "Your hair looks great!" You can also give high fives, or bones when they do something they should be doing. Even if your tutee does not respond, it does not mean they are not

listening. I know that I like to hear when I do something right and I am sure you do as well. The same is true with your tutee. Keep trying, you may be surprised at how your tutee reacts to your kind words."

APPENDIX C

PEER TUTOR EVALUATION – MAIN TRAINING SESSION

Peer Tutor Evaluation – Main Training Session

Name:_____

Age:_____

1. What does a tutor do in Special Physical Education Class?

a. help students learn new skills
b. tell students what they are doing well
c. tell students what they can do to improve
d. be friendly, smile, and talk to the students
e. all of the above

2. How did you like the Peer Tutor Training?

Circle one:

 

3. How much fun did you have in the Peer Tutor Training?

Circle a number: 1 = no fun at all, 10 = the most fun possible

1 2 3 4 5 6 7 8 9 10

4. What type of feedback should you use the most?

a. negative feedback
b. positive feedback

5. How should you talk to your tutee?

a. loudly
b. look at your tutee and talk quietly

6. What cues do you give to your tutee when they are practicing kicking a ball?

a. step, tick, kick
b. kick, step, tick

APPENDIX D

3-POINT LIKERT SCALE ATTITUDE INSTRUMENT

3-Point Peer Tutor Likert Attitude Instrument

Date:	Tutor	Attitude		
				
				
				
				
				
				
				
				
				
				
				
				

REFERENCES

Ajzen, L., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice Hall.

Allport, G.W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.

Antonak, R.F., & Livneth, H. (1988). *The measurement of attitudes toward people with disabilities*. Springfield, IL: Charles C. Thomas.

Barfield, J., Hannigan-Downs, S., & Lieberman, L. (1998). Implementing a peer tutor program: Strategies for practitioners. *The Physical Educator, 55*(4), 211-21.

Block, M. E. (1994). Why all students with disabilities should be included in regular physical education. *Palaestra, 10*(3), 17-24.

Block, M. E. (1995). Development and validation of the Children's Attitudes toward Integrated Physical Education-Revised (CAIPE-R) inventory. *Adapted Physical Activity Quarterly*, *12*(1), 60-77.

Block, M. E. (2007). *Teacher's guide to including students with disabilities in general physical education*. Baltimore, MD: Paul H. Brookes Publishing Co.

Block, M., Oberweiser, B., & Bain, M. (1995). Using classwide peer tutoring to facilitate inclusion of students with disabilities in regular physical education. *The Physical Educator, 52*, 47-56.

Bogdan, R., & Biklen, S. K. (1992). Qualitative research: An introduction to theory and methods. Boston, MA: Allyn & Bacon.

Bourke, S., & Burgman, I. (2010). Coping with bullying in Australian schools: How children with disabilities experience support from friends, parents and teachers. *Disability & Society*, *25*(3), 359-371.

Breslin, C. M., & Rudisill, M. E. (2013). Relationships among assessment time, time on task, and motor skill performance in children with autism spectrum disorder. *Adapted Physical Activity Quarterly, 30*(4), 338-350.

Carr, J. E. (2005). Recommendations for reporting multiple-baseline designs across participants. *Behavioral Interventions, 20*(3), 219-224. doi:10.1002/bin.191

Cervantes, C. M., Lieberman, L. J., Magnesio, B., & Wood, J. (2013). Peer tutoring. *The Journal of Physical Education, Recreation & Dance, 84*(3), 43-48.

Charlop, M.H., Schreibman, L., & Tyron, A. (1983). Learning through observation: The effects of peer modeling on acquisition and generalization in autistic children. *Journal of Abnormal Child Psychology*, *11*, 355-365.

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine*, *119*(2), 166-e7.

d'Arripe-Longueville, F., Gernigon, C., Huet, M., Cadopi, M., & Winnykamen, F. (2002). Peer tutoring in a physical education setting: Influence of tutor skill level on novice learners' motivation and performance. *Journal of Teaching in Physical Education, 22*, 105-123.

d'Arripe-Longueville, F., Gernigon, C., Huet, M. L., Winnykamen, F., & Cadopi, M. (2002). Peer assisted learning in the physical activity domain: Dyad type and gender differences. *Journal of Sport & Exercise Psychology, 24*, 219-238.

de Boer, A., Pijl, S. J., & Minnaert, A. (2012). Students' attitudes towards peers with disabilities: A review of the literature. *International Journal of Disability, Development and Education*, *59*(4), 379-392.

De Lisi, R. (2002). From marbles to instant messenger: Implications of Piaget's ideas about peer learning. *Theory into Practice, 41*, 5-12.

Drew, C. J., Hardman, M. L., & Hosp, J. L. (2007). *Designing and conducting research in education*. Thousand Oaks, CA: Sage Publications.

Dunn, J. M., & Leitschuh, C. A. (2010). *Special physical education* (9th ed.). Dubuque, IA: Kendall/Hunt.

Ensergueix, P., & Lafont, L. (2010). Reciprocal peer tutoring in a physical education setting: Influence of peer tutor training and gender on motor performance and self-efficacy outcomes. *European Journal of Psychology of Education, 25*(2), 222-242. doi: 10.1007/s10212-009-0010-0

Fenrick, N., & McDonnell, J. (1980). Junior high students as teachers of the severely retarded: Training and generalization. *Education and Training of the Mentally Retarded, 15,* 187-194.

Field, A. P. (2013). *Discovering statistics using SPSS* (4th ed.). London, England: SAGE.

Greenwood, C.R., & Delquardri, J., (1991). Code for instructional structure and student academic response (CISSAR). In M. Herson & A. S. Bellack (Eds.), *Dictionary of*

behavioral assessment (pp. 120-122). New York: Pergamon.

Henderson, H., Lavay, B., & French, R. (2006). *Positive behavior management in physical activity settings* (2nd ed.). Champaign, IL: Human Kinetics.

Hodge, S. R., LaMaster, K., Murata N. M., Casebolt, K., O'Sullivan, M., & Ammah, J. (2000). *Analysis of inclusion practices in physical education*. Unpublished systematic observation instrument, The Ohio State University, Columbus, OH.

Hodge, S. R., Lieberman, L. J., & Murata, N. M. (2012). *Essentials of teaching adapted physical education: Diversity, culture, and inclusion*. Scottsdale, AZ: Holcomb Hathaway, Incorporated.

Hogan, D.M., & Tudge, J.R. (1999). Implications of Vygotsky's theory for peer learning. In A.M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning* (pp. 39-65). Mahwah, NJ: Lawrence Erlbaum.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, *71*(2), 165-179.

Houston-Wilson, C., Dunn, J. M., van der Mars, H., & McCubbin, J. (1997). The effect of peer tutors on motor performance in integrated physical education classes. *Adapted Physical Activity Quarterly, 14*(4), 298-313.

Houston-Wilson, C., Lieberman, L., Horton, M., & Kasser, S. (1997). Peer tutoring: A plan for instructing students of all abilities. *Journal of Physical Education, Recreation & Dance,* 68(6), 39-44.

Johnson, M., & Ward, P. (2001). Effects of class-wide peer tutoring on correct performance of striking skills in 3rd grade physical education. *Journal of Teaching in Physical Education, 20,* 247-263.

Kamps, D. M., Barbetta, P. M., Leonard, B. R., & Delquadri, J. (1994). Classwide peer tutoring: An integration srategy to improve reading skills and promote peer interactions among students with autism and general education peers. *Journal of Applied Behavior Analysis, 27,* 49-62.

Kazdin, A. E. (1982). *Single case research design: Methods for clinical and applied settings*. New York: Oxford University Press.

Kendall, M., & Gibbsons, D.J. (1990). *Rank correlation methods* (5th ed.). London: Edward Arnold.

Klavina, A. (2008). Using peer-mediated instructions for students with severe and multiple disabilities in inclusive physical education: A multiple case study. *European*

Journal of Adapted Physical Activity, 1(2), 7-19.

Klavina, A., & Block, M. E. (2008). The effect of peer tutoring on interaction behaviors in inclusive physical education. *Adapted Physical Activity Quarterly, 25*(2), 132-158.

Knowles, C. J., Aufderheide, S. K., & McKenzie, T. (1982). Relationship of individualized teaching strategies to academic learning time for mainstreamed handicapped and nonhandicapped students. *The Journal of Special Education*, *16*(4), 449-456. doi:10.1177/002246698201600408.

Kunin, T. (1955). The construction of a new type of attitude measure. *Personnel Psychology*, *8*(1), 65-77.

Lancioni, G.E. (1982). Normal children as tutors to teach social responses to withdrawn mentally retarded school mates: Training, maintenance, and generalization. *Journal of Applied Behavior Analysis, 15*, 17-40.

Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. Cambridge, UK: Cambridge University Press.

Lieberman, L. J., Dunn, J. M., van der Mars, H., & McCubbin, J. (2000). Peer tutors' effects on activity levels of deaf students in inclusive elementary physical education. *Adapted Physical Activity Quarterly, 17*(1), 20-39.

Lieberman, L. J., & Houston-Wilson, C. (2012). *Strategies for inclusion: A handbook for physical educators*. Champaign, IL: Human Kinetics.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22*(140), 55.

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, *35*(6), 382-386.

Macklin, M. C., & Machleit, K. A. (1990). Measuring preschool children's attitude. *Marketing Letters*, (*13*), 253-265.

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D.X. (1985). Teacher judgements concerning acceptability of school-based interventions. *Professional Psychology: Research and Practice, 16,* 191-198.

Martens, B. K., Witt, J. C., Elliott, S. N., Darveaux, D., & Lane, K. L. (1999). Intervention Rating Profile--15. *Journal of Emotional and Behavioral Disorders*, (*7)* 211-223.

McDonnell, J., Thorson, N., & Allen, C. (2000). The effects of partner learning during spelling for students with severe disabilities and their peers. *Journal of Behavioral*

Education, *10*(23), 107-121.

O'Neill, R. E., McDonnell, J. J., & Billingsley, F. F. (2011). *Single case research designs in educational and community settings*. New York: Pearson Education, Limited.

Owen-DeSchryver, J. S., Carr, E. G., Cale, S. I., & Blakeley-Smith, A. (2008). Promoting social interactions between students with autism spectrum disorders and their peers in inclusive school settings. *Focus on Autism & Other Developmental Disabilities*, *23*(1), 15-28. doi: 10.1177/1088357608314370

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem in intellectual development*. Chicago, IL: University of Chicago Press.

Piatti, C., Beets, A., & Combs, C. (2010). In-school physical activity for children with intellectual disabilities. *JOPERD: The Journal of Physical Education, Recreation & Dance*, *81*(7), 11.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Rosenshine, B. (1979). Content time and direct instruction. In P.L. Peterson & H.J. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications* (pp. 28-53). Berkeley, CA: McCutchan.

Şad, S. N. (2012). An attitude scale for smart board use in education: Validity and reliability studies. *Computers & Education*, *58*(3), 900-907.

Siedentop, D., & Tannehill, D. (2000). *Developing teaching skills in physical education*. Mountain View, CA: Mayfield.

Siegel, S., & Castellan, Jr., N.J. (1988). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Silverman, S., Dodds, P., Placek, J., Shute, S., & Rife, F. (1984). Academic Learning Time in Elementary School Physical Education (ALT-PE) for student subgroups and instructional activity units. *Research Quarterly For Exercise & Sport*, *55*(4), 365-370.

Shadish, W., & Cook, T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Sherrill, C. (1998). *Adapted physical activity recreation and sport: Crossdisciplinary and lifespan* (5th ed.). Dubuque, IA: WCB/McGraw-Hill.

Skinner, B.F. (1974). *About behaviorism*. London: Jonathan Cape.

Slininger, D., Sherrill, C., & Jankowski, C. M. (2000). Children's attitudes toward peers

with severe disabilities: Revisiting contact theory. *Adapted Physical Activity Quarterly*, *17*(2), 176-196.

Snyder, T. D., & Dillow, S. A. (2011). *Digest of education statistics 2010 (NCES 2011-015)*. Washington DC; National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Thomas, J. R., Silverman, S., & Nelson, J. (2015). *Research Methods in Physical Activity, 7E*. Champaign, IL: Human Kinetics.

Triandis, H. (1971). *Attitude and attitude change*. New York: Wiley & Sons.

Tripp, A., French, R., & Sherrill, C. (1995). Contact theory and attitudes of children in physical education programs toward peers with disabilities. *Adapted Physical Activity Quarterly*, *12*, 323-323.

Ulrich, D. A. (2000). *Test of gross motor development-2*. Austin, TX: Pro-Ed.

Vygotsky, L.S. (1978) *Mind in society. The development of higher psychological processes* (A. R. Luria, M. Lopez-Morillas, M., Cole & J. Wertsch, Trans). Cambridge, MA: Harvard University Press.

Ward, P., & Ayvazo, S. (2006). Classwide peer tutoring in physical education: Assessing its effects with kindergartners with autism. *Adapted Physical Activity Quarterly*, *23*(3), 233.

Ward, P., & Lee, M. (2005). Peer-assisted learning in physical education: A review of theory and research. *Journal of Teaching in Physical Education, 24*(3), 205.

Webster, G. E. (1987). Influence of peer tutors upon academic learning time-physical education of mentally handicapped students. *Journal of Teaching in Physical Education, 6*(4), 393-403.

Welk, G. (2002). *Physical activity assessments for health-related research*. Champaign, IL: Human Kinetics.

Wessel, J. A., & Zittel, L. L. (1998). *I Can Primary Skills K-3*. Austin, TX: Pro-Ed.

Wiskochil, B., Lieberman, L. J., Houston-Wilson, C., & Petersen, S. (2007). The effects of trained peer tutors on the physical education of children who are visually impaired. *Journal of Visual Impairment & Blindness, 101*(6), 339-350.

Witt, J.C., & Martens, B.K. (1983). Assessing the acceptability of behavioral interventions used in classrooms. *Psychology in the Schools, 20,* 510-517.

Wolery, M., & Harris, S. R. (1982). Interpreting results of single-subject research

designs. *Physical Therapy, 62*, 445-452.

Yun, J., & Ulrich, D. A. (2002). Estimating measurement validity: A tutorial. *Adapted Physical Activity Quarterly, 19*, 32-47.